

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 4 年 3 月 1 0 日
Date of Application:

出 願 番 号 特 願 2 0 0 4 - 0 6 6 5 4 8
Application Number:
[ST. 10/C]: [J P 2 0 0 4 - 0 6 6 5 4 8]

出 願 人 株式会社日立製作所
Applicant(s):

CERTIFIED COPY OF
PRIORITY DOCUMENT

2 0 0 4 年 3 月 3 1 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 4 - 3 0 2 6 4 0 2

【書類名】 特許願
【整理番号】 K04004601A
【あて先】 特許庁長官殿
【国際特許分類】 G06F 3/06
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 中川 弘隆
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 山本 政行
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 兼田 泰典
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社 日立製作所
【代理人】
 【識別番号】 100075096
 【弁理士】
 【氏名又は名称】 作田 康夫
【選任した代理人】
 【識別番号】 100100310
 【弁理士】
 【氏名又は名称】 井上 学
【先の出願に基づく優先権主張】
 【出願番号】 特願2004- 6214
 【出願日】 平成16年 1月14日
【手数料の表示】
 【予納台帳番号】 013088
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1
 【包括委任状番号】 9902691
 【包括委任状番号】 0315335

【書類名】 特許請求の範囲**【請求項 1】**

計算機とネットワークを介して接続し前記計算機が利用するデータを格納する記憶領域を有する前記記憶装置にネットワークを介して接続する管理計算機で、複数の記憶装置の記憶領域を管理する方法であって、
記憶領域の所定の性能を示すレベルと前記記憶装置の装置特性情報との対応付けを保持し

、
第 1 の記憶装置が前記計算機に割り当てた記憶領域の性能を示すレベルを第 1 の記憶装置から取得し、

前記取得した記憶領域の性能を示すレベルに対応する前記第 1 の装置特性情報と、他の記憶装置の装置特性情報とをそれぞれ参照し、記憶装置間での前記記憶領域の性能を比較することを特徴とする記憶領域管理方法。

【請求項 2】

前記対応付けは、前記計算機と接続する記憶装置から取得したものであることを特徴とする請求項 1 記載の記憶領域管理方法。

【請求項 3】

前記記憶装置間での前記記憶領域の性能の比較は、前記管理計算機と前記ネットワークを介して接続した新たな記憶装置から前記対応付けを取得した場合に行うことを特徴とする請求項 1 記載の記憶領域の管理方法。

【請求項 4】

前記記憶装置間での前記記憶領域の性能の比較は、前記記憶装置から新たな対応付けをさらに取得した場合に行うことを特徴とする請求項 2 記載の記憶領域管理方法。

【請求項 5】

前記レベルは、0 から 10 までの整数値であることを特徴とする請求項 1 記載の記憶領域管理方法。

【請求項 6】

前記比較結果に基づいて前記他の記憶装置に対し、前記取得したレベルに対応する前記他の記憶装置の装置特性情報を備える記憶領域を前記計算機に割り当てるよう指示すること
特徴とする請求項 1 記載の記憶領域管理方法。

【請求項 7】

前記管理計算機は、前記取得したレベルに対応する前記他の記憶装置の装置特性情報を備える記憶領域を計算機に割り当てる旨の入力を受け、前記割り当ての指示を行うことを特徴とする請求項 6 記載の記憶領域管理方法。

【請求項 8】

記憶領域の所定の性能を示すレベルが前記所定の性能を有する必要があることを示す場合は、前記所定の性能を示すレベルについて前記比較を行わないことを特徴とする請求項 1 記載の記憶領域管理方法。

【請求項 9】

前記計算機と前記記憶装置とは第 1 のネットワークを介して接続し、前記管理計算機は、前記計算機及び前記記憶装置に前記第 2 のネットワークを介して接続し、前記指示に基づいて前記計算機に割り当てる記憶領域に、前記第 1 の記憶装置が有する記憶領域に格納されているデータを複写し、

前記計算機が前記割り当てる記憶領域に対し前記複写したデータの読み出し及び新たなデータの書き込みの少なくとも一方を前記第 1 のネットワークを介して実行するように前記計算機に前記第 2 のネットワークを介して指示することを特徴とする請求項 6 記載の記憶領域管理方法。

【請求項 10】

所定の性能が複数有する場合、前記比較は、そのレベルが一番高いレベルで行う特徴とする請求項 1 記載の記憶領域管理方法。

【請求項 11】

計算機とネットワークを介して接続する第 1 の記憶装置と、前記第 1 の記憶装置と接続する第 2 の記憶装置とにネットワークを介して接続する管理計算機で、前記第 1 及び第 2 の記憶装置がそれぞれ有する記憶領域を管理する方法であって、記憶領域の所定の性能を示すレベルと記憶装置の装置特性情報との対応付けを前記記憶装置ごとに保持し、

第 1 の記憶装置が有する記憶領域の所定の性能を示すレベルと、当該第 1 の記憶装置が有する記憶領域と接続する第 2 の記憶装置が有する記憶領域の所定の性能を示すレベルとを取得し、

前記取得したレベルに対応する装置特性情報を比較し、

前記比較結果に基づいて前記第 1 の記憶装置が有する記憶領域に前記第 2 の記憶装置が有する記憶領域に格納されているデータを格納することを特徴とする記憶領域管理方法。

【請求項 1 2】

前記第 1 の記憶装置が有する記憶領域が前記第 2 の記憶装置が有する記憶領域と接続することを示す階層情報を取得し、前記階層情報に基づいて前記第 1 及び第 2 の記憶装置それぞれ有する前記装置特性情報を比較することを特徴とする請求項 1 1 記載の記憶領域管理方法。

【請求項 1 3】

前記階層情報は、前記第 1 の記憶装置から取得することを特徴とする請求項 1 2 記載の記憶領域管理方法。

【請求項 1 4】

前記比較結果、前記第 2 の記憶装置が有する記憶領域の所定の性能を示すレベルに対応する前記第 1 の記憶装置の装置特性情報が、前記第 2 の記憶装置の装置特性情報より良い場合は、前記第 1 の記憶装置に前記レベルに対応する前記第 1 の記憶装置の装置特性情報に基づいて所定の性能を備える記憶領域に前記データを格納することを特徴とする請求項 1 2 または 1 3 いずれか記載の記憶領域管理方法。

【請求項 1 5】

前記階層情報を消去するよう前記第 1 の記憶装置に指示することを特徴とする請求項 1 4 記載の記憶領域管理方法。

【請求項 1 6】

計算機とネットワークを介して接続する第 1 の記憶装置であって、

前記計算機が利用するデータを格納する他の記憶装置が有する記憶領域と接続する記憶領域と、

前記記憶領域の所定の性能を示すレベルと前記記憶装置の装置特性情報との記憶装置ごとの対応付けを保持するメモリと、

前記計算機から前記第 1 の記憶装置又は前記他の記憶装置に対するアクセスを制御する制御部とを備え、

前記制御部は、

前記他の記憶装置が有する記憶領域の所定の性能を示すレベルとを取得し、前記レベルに対応する前記第 1 及び前記他の記憶装置の装置特性情報を前記対応付けから参照し、前記参照した値を比較することを特徴とする第 1 の記憶装置。

【請求項 1 7】

前記比較結果に基づいて、前記第 1 の記憶装置が有する記憶領域の所定の性能を示すレベルに対応する装置特性情報を備える記憶領域に前記データを格納することを特徴とする請求項 1 6 記載の第 1 の記憶装置。

【請求項 1 8】

計算機が利用するデータを格納する記憶領域を有する複数の記憶装置と、前記記憶装置とネットワークを介して接続する管理計算機とを有する計算機システムであって、

前記管理計算機は、前記記憶装置から受ける情報に基づいて処理を行う制御部と、前記記憶装置への要求及び前記記憶装置からの応答を行う管理ネットワークインタフェースとを有し、

前記記憶装置各々は、前記記憶領域の所定の性能を示すレベルと、前記記憶装置の前記性能の値とを対応付けと前記記憶装置が有する記憶領域の性能を示すレベルとを保持するメモリと、前記対応付けを参照し、前記管理計算機からの要求する所定の性能を示すレベルに対応する性能の値を備える前記記憶領域を前記計算機に割り当てる記憶領域割当部と、前記管理計算機とのデータの授受を行うネットワークインタフェースとを有し、前記管理計算機の制御部は、前記管理ネットワークインタフェースに前記ネットワークを介して接続する前記ネットワークインタフェースに接続されている記憶装置を検出した場合、前記記憶装置から前記対応付けを取得し、前記記憶領域制御モジュールが前記計算機へ既に割り当てられた記憶領域の性能を示すレベルを取得し、前記複数の記憶装置の対応付けを参照し、前記レベルに対応する装置特性情報を前記複数の記憶装置同士で比較し、前記比較結果に基づいて前記計算機に割り当てる記憶領域を既に割り当てられた記憶領域を備える記憶装置とは異なる記憶装置へ前記レベルに対応する記憶領域の割り当てを指示し、

前記記憶装置の記憶領域割当部は、前記割り当て指示に基づいて前記レベルに対応する性能を備える記憶領域を前記計算機へ割り当てることを特徴とする計算機システム。

【請求項 19】

計算機とネットワークを介して接続し前記計算機が利用するデータを格納する記憶領域を有する前記記憶装置にネットワークを介して接続する管理計算機で実行される複数の記憶装置の記憶領域を管理する管理プログラムであって、記憶領域の所定の性能を示すレベルと前記記憶装置の装置特性情報との対応付けを取得する手順と、

第1の記憶装置が前記計算機に割り当てた記憶領域の性能を示すレベルを第1の記憶装置から取得する手順と、

前記取得した記憶領域の性能を示すレベルに対応する前記第1の装置特性情報と、他の記憶装置の装置特性情報とをそれぞれ参照し、記憶装置間での前記記憶領域の性能を比較する手順とを前記管理計算機で実行させることを特徴とする管理プログラム。

【請求項 20】

請求項 19 記載の管理プログラムをコンピュータ読み取り可能に記憶した記憶媒体。

【書類名】明細書

【発明の名称】記憶領域の管理方法、管理装置及び管理プログラム

【技術分野】

【0001】

本発明は複数の記憶装置を含む計算機システムにおける記憶領域の管理方法に関する。

【背景技術】

【0002】

近年、複数のサーバ、複数の記憶装置、バックアップ装置等を、スイッチやハブ等で接続するストレージネットワーク構成（主にSAN（Storage Area Network）と呼ばれる）が用いられている。ストレージネットワークによって計算機システムを構成するメリットは、スケーラビリティに優れていること、複数の記憶装置に分散したデータを統合して一元管理を行えるため管理コストの低減を実現できること等があげられる。

【0003】

また、ディスクアレイ装置に代表される大規模な記憶装置においては、記憶装置内に計算機が利用する情報（データやプログラム）を格納するための記憶領域を作成するための管理手段が用意されている。しかし、記憶領域を作成するための管理手段は、記憶装置を製造するベンダごとにそのインタフェースや要求のフォーマットが異なっている。このため、異なるベンダの記憶装置を導入した場合には、記憶装置の管理者は、ベンダごとに異なるインタフェースや要求のフォーマットを覚えなければならず、大規模なコンピュータシステムでは、大きな問題となっている。そこで、DMTF（Distributed Management Task Force）は、CIM（Common Information Model）とWBEM（Web-Based Enterprise Management）を策定し、記憶装置において記憶領域を作成する際のインタフェースや要求のフォーマットを策定している。このCIMで定義されたインタフェースを用いることで、ベンダの異なる記憶装置においても、統一的な手法で記憶領域を作成することができる。このCIMで定義された記憶領域の作成インタフェースと要求のフォーマットでは、「ヒント」を用いた記憶領域の作成が定義されている。

(Device27_StorageServices.mofより抜粋)

```
// =====  
// StorageSettingWithHints  
// =====  
[Experimental, Version("2.7.1"), Description (  
    "This subclass of StorageSetting allows a client to specify "  
    "'hint's for optimization of the volume performance. The effect "  
    "of these hints is implementation dependent.") ]  
class CIM_StorageSettingWithHints: CIM_StorageSetting {  
  
    [MinValue (0), MaxValue (10), Description (  
        "This hint is an indication from a client of the importance "  
        "placed on data availability. Values are 0=Don't Care to "  
        "10=Very Important.") ]  
    uint16 DataAvailabilityHint;  
  
    [MinValue (0), MaxValue (10), Description (  
        "This hint is an indication from a client of the randomness "  
        "of accesses. Values are 0=Entirely Sequential to "  
        "10=Entirely Random.") ]  
    uint16 AccessRandomnessHint;  
  
    [MinValue (0), MaxValue (10), Description (  

```

```

    "This hint is an indication from a client of the direction "
    "of accesses. Values are 0=Entirely Read to "
    "10=Entirely Write.") ]
uint16 AccessDirectionHint;

[Description (
    "This hint is an indication from a client of the optimal "
    "access sizes. Several sizes can be specified."),
    Units ("MegaBytes") ]
uint16 AccessSizeHint[];

[MinValue (0), MaxValue (10), Description (
    "This hint is an indication from a client how important "
    "access latency is. Values are 0=Don't Care to "
    "10=Very Important.") ]
uint16 AccessLatencyHint;

[MinValue (0), MaxValue (10), Description (
    "This hint is an indication from a client of bandwidth "
    "prioritization. Values are 0=Don't Care to "
    "10=Very Important.") ]
uint16 AccessBandwidthWeight;

[MinValue (0), MaxValue (10), Description (
    "This hint is an indication of the importance the client "
    "places on the cost of storage. Values are 0=Don't Care to "
    "10=Very Important. A StorageVolume provider might choose "
    "to place data on low cost or high cost drives based on "
    "this parameter.") ]
uint16 StorageCostHint;

[MinValue (0), MaxValue (10), Description (
    "This hint is an indication of the importance placed on "
    "storage efficiency by the client. Values are 0=Don't Care "
    "to 10=Very Important. A StorageVolume provider might choose "
    "different RAID levels based on this hint.") ]
uint16 StorageEfficiencyHint;
};

```

「ヒント」を用いることで、作成する記憶領域を抽象的に指定することが出来る。このため、各ベンダの記憶装置のハードウェア構成がまったく異なっても、統一的な要求のフォーマットによる記憶領域の作成を提供することができる。

【 0 0 0 4 】

一方、特開 2 0 0 3 - 1 4 0 8 3 6 には、S A N 環境における複数の記憶装置において、管理者が記憶領域の再配置による配置最適化を簡便に行うための方法が開示されている。その概要は次の通りである。S A N 環境に接続されている複数の記憶装置を構成する記憶装置を複数の組（クラス）に分類、または前記記憶装置を構成する複数のディスク装置を複数の組（クラス）に分類し、そのクラスにクラス属性（使用状況の閾値）を設定する。S A N に接続された S A N 管理サーバおよび前記記憶装置の記憶領域を利用するサーバのいずれか一方は、前記複数の記憶装置から、前記記憶装置のベンダ情報（製造元情報）と使用状況情報および前記クラス属性を取得して、前記ベンダ情報（製造元情報）、前記

使用状況情報および前記クラス属性情報に基づき、第一の前記クラス内にある記憶領域の再配置先に適する第二の前記クラス内にある記憶領域を選択し、前記第一の記憶領域のデータを前記第二の記憶領域へコピーするとともに記憶領域の対応づけを前記第一の記憶領域から前記第二の記憶領域へ変更して再配置を行うことを特徴とする記憶装置の管理方法である。

【0005】

【特許文献1】特開2003-140836号公報

【非特許文献1】Device27_StorageServices.mof (662～721行目)、[online]、2002年11月12日、Distributed Management Task Force、[2003年1月18日検索]、インターネット<URL: http://www.dmtf.org/standards/cim_schema_v27.php>

【発明の開示】**【発明が解決しようとする課題】****【0006】**

特開2003-240836号公報に開示されている記憶領域の再配置方法によれば、クラス単位の使用状況情報、クラス属性および記憶装置のベンダ情報により、クラス間またはクラス内の記憶領域の移行を行うことが開示されている。

【0007】

しかし、ネットワークで接続された複数の記憶装置間での記憶領域の移行をおこなうことについては、開示されていない。また、ヒントは、各ベンダの記憶装置のハードウェア構成がまったく異なっている、統一的な要求のフォーマットによる記憶領域の作成を提供することができるが、ヒントが同じ値であっても、それに対応する実際の装置の特性は、各ベンダなどの記憶装置間では異なる場合が多い。したがって、複数の記憶装置間で記憶領域を行う者がヒントによって複数の記憶装置間で、リソースを有効利用するための記憶領域割り当てを実現することは困難である。

【0008】

本発明は、上記点を鑑みて考えられたものであり、記憶領域を割り当てる際に指定する記憶領域の所定の性能を示すレベル（ヒント）を用いて複数の記憶装置のリソース（記憶領域）の有効利用を実現することを目的とする。

【課題を解決するための手段】**【0009】**

本発明の目的を達成するための一つの実施形態として、計算機とネットワークを介して接続し計算機が利用するデータを格納する記憶領域を有する前記記憶装置にネットワークを介して接続する管理計算機は、記憶領域の所定の性能を示すレベルと前記記憶装置の装置特性情報との対応付けを保持し、第1の記憶装置が前記計算機に割り当てた記憶領域の性能を示すレベルを第1の記憶装置から取得し、取得した記憶領域の性能を示すレベルに対応する前記第1の装置特性情報と、他の記憶装置の装置特性情報とをそれぞれ参照し、記憶装置間での前記記憶領域の性能を比較することを特徴とする。

【0010】

なお、本発明の他の特徴は、本明細書及び添付図面の記載により明らかにする。

【発明の効果】**【0011】**

ネットワークを介する複数の記憶装置のリソースを有効に利用することができる。

【発明を実施するための最良の形態】**【0012】**

以下、本発明の実施の形態を図面を参照しながら詳細に説明する。

●実施形態の概要

図1を用いて本発明を実施するための形態の概要を説明する。図1の計算機システムでは、記憶領域を利用する計算機100と記憶装置301および記憶装置302を第一のネットワーク200で接続し、記憶装置301および記憶装置302を管理する管理計算機

600と記憶装置301と302を第二のネットワーク500で接続している。図1のシステム構成は、計算機100と記憶装置301および管理計算機600からなる計算機システムに、記憶装置302が追加（構成変更）した状態である。また、第一のネットワーク200には、記憶領域間でデータの移動を行う記憶領域移動実行手段700を設けている。記憶領域移動実行手段700は、管理計算機600からの移動要求を受け付けられるように、第二のネットワークにも接続している。

【0013】

管理計算機600は、記憶装置302が追加されたことを検出すると（10）、記憶装置301と記憶装置302のそれぞれから、装置特性情報と「ヒント」を取得し（12）、装置特性情報とヒントを再評価し、記憶領域を新しく追加した記憶装置302に設けた方が、よりヒントに合致した記憶領域を提供できると判断した場合には、当該記憶領域に関連付けられた「ヒント」によって、記憶装置302に記憶領域を作成し（14）、記憶装置301の記憶領域のデータを記憶領域移動実行手段700を用いて、記憶装置302に新たに作成した記憶領域に移行する。最後に管理計算機600は、計算機100に記憶領域を移動したことを通知し、計算機100は、それ以降記憶装置302の記憶領域を参照するように設定することで（18）、計算機に対し、より「ヒント」に合致した記憶領域を使用できるようにする。

【実施例1】

【0014】

（1）システム構成

図2に実施例1のシステム構成を示す。図2の計算機システムでは、記憶領域を利用する計算機100と記憶装置301および記憶装置302を第一のネットワーク200で接続し、記憶装置301および記憶装置302を管理する管理計算機600と記憶装置301と302を第二のネットワーク500で接続している。図2のシステム構成は、計算機100と記憶装置301および管理計算機600からなる計算機システムに、記憶装置302を追加（構成変更）した状態である。また、第一のネットワーク200には、記憶領域間でデータの移動を行う記憶領域移動実行手段700を設けている。記憶領域移動実行手段700は、管理計算機600からのデータ移動要求を受け付けられるように、第二のネットワークにも接続している。本実施例では、第一のネットワーク200としては、ファイバチャネルを用いており、また、第二のネットワーク500としてはIP（Internet Protocol）ネットワークを用いている。第1及び第2のネットワークは上記に限られない。また、第1及び第2のネットワークは、異なるネットワークであっても同じネットワークであってもよい。

●計算機

計算機100は、プログラムの実行を司るCPU105と、プログラムやプログラムの実行に必要なデータを格納するメモリ110と、第一ネットワーク200へ接続し記憶装置301および記憶装置302とデータの交換を行うためのファイバチャネルインタフェース（以下FCインタフェースと略す）120と、第二のネットワークに接続するためのネットワークインタフェース130からなる。計算機100は、プログラムの実行によって生成したデータを、FCインタフェース120を介して第一のネットワーク200に接続された記憶装置の記憶領域に格納することができる。また、計算機100は、プログラムそのものやプログラムの実行に必要なデータを、FCインタフェース120を介して記憶装置の記憶領域から取得することができる。

【0015】

そして、計算機100のメモリ110にはボリューム管理テーブル115を設けている（図6）。ボリューム管理テーブル115は、計算機100が計算機上のアプリケーションからファイルシステムを経由して利用するボリュームが、どの記憶装置のどこ記憶領域であるかを管理するためのテーブルである。本実施例では、このボリューム管理テーブル115をボリュームに割り当てられたドライブレターと、FCインタフェースの番号、記憶領域の番号で識別する。FCインタフェースの番号としては、WWN（World Wide Nam

e) を用いることが一般的である。また、記憶領域の番号は、S C S I (Small Computer System Interface) で規定される L U N (Logical Unit Number) を用いることが一般的である。本実施例では、図との対応を簡易化するために、WWN と L U N に変えて、図で使用する番号をそのまま用いて説明する。計算機 1 0 0 は、このボリューム管理テーブル 1 1 5 にしたがって、記憶領域からのデータの読み出しや記憶領域へのデータの書き込みを行う。本実施例では、管理計算機 6 0 0 からネットワークインタフェース 1 3 0 を介して、ボリューム管理テーブル 1 1 5 を書き換えることができる構成としている。

●記憶装置

記憶装置の構成を説明する。記憶装置 3 0 1 と 3 0 2 は、プログラムの実行を司る C P U 3 5 0 と、プログラムやプログラムの実行に必要な情報を格納するメモリ 3 6 0 と、管理計算機 6 0 0 との通信を行うためのネットワークインタフェース 3 9 0 と、計算機 1 0 0 とデータの交換を行うための F C インタフェース (3 7 1 1 ~ 3 7 1 6, 3 7 2 1 ~ 3 7 2 9) と、計算機 1 0 0 からのデータのリード/ライト要求を処理する記憶領域制御モジュール 3 8 0 と、計算機 1 0 0 から受け取ったデータや記憶領域から読み出したデータを一時的に蓄えるためのキャッシュ 3 8 9 と、実際にデータを蓄積するための記憶領域とを有する。

記憶装置では、計算機に割り当て済みの記憶領域を記憶領域制御モジュール 3 8 0 が管理する記憶領域情報によって割当済みの記憶領域と未割当の記憶領域を管理している。本実施例では記憶装置 3 0 1 は、記憶領域情報 3 8 3 を有し、記憶装置 3 0 2 は、記憶領域情報 3 8 4 を有する。

【 0 0 1 6 】

メモリ 3 6 0 は、装置特性情報 3 2 5 と、記憶領域情報 3 8 2 とを保持し管理計算機 6 0 0 からの記憶領域割当要求にしたがって、記憶領域を割り当てる記憶領域割当プログラム 3 3 0 を有している。なお、プログラムは、記憶装置内の ROM や磁気ディスク等の不揮発性記憶媒体に記憶され、記憶装置の起動時にメモリ 3 6 0 にロードされて実行されるものとする。また、プログラムの処理を記憶装置内にハードウェア的に構成してもよい。

【 0 0 1 7 】

記憶装置 3 0 1 と 3 0 2 の違いは、F C インタフェースの提供できる帯域と数である。記憶装置 3 0 1 には、6 個の F C インタフェースを有し、帯域が 1 G b p s の F C インタフェース 3 7 1 1 ~ 3 7 1 3 と、帯域が 2 G b p s の F C インタフェース 3 7 1 4 ~ 3 7 1 6 を有している。一方、記憶装置 3 0 2 には、9 個の F C インタフェースを有し、帯域が 1 G b p s の F C インタフェース 3 7 2 1 ~ 3 7 2 3 と、帯域が 2 G b p s の F C インタフェース 3 7 2 4 ~ 3 7 2 6 と、帯域が 1 0 G b p s の F C インタフェース 3 7 2 7 ~ 3 7 2 9 を有している。

【 0 0 1 8 】

図 3 は、管理計算機 6 0 0 から、記憶装置に発行される記憶領域割当要求に含まれるパラメータの例である。本実施例では、パラメータとして、作成する記憶領域の容量と、ホストからのアクセス帯域に関して、広帯域アクセスを希望する度合いを示す AccessBandwidthHint (ABH)、記憶装置内の記憶領域のビット単価について、ビット単価の安い記憶領域を希望する度合いを示す StorageCostHint (SCH)、記憶領域の可用性に関して、高い可用性を持つ記憶領域を希望する度合いを示す DataAvailabilityHint (DAH) の四つのヒントのうち、三つのヒント、ABH、SCH、DAH を用いて説明する。それぞれのヒントには 1 ~ 1 0 の値により 1 0 段階のレベルを与えることができる。このレベルは、それぞれのヒントが意味する記憶領域の性能を重要視する度合いを示す。例えばレベル 0 は「Don't care」であり、当該ヒントを重要視しないことを示し、そのヒントに該当する記憶領域の性能は記憶装置各々で低い性能となる。そして、レベルの数字が大きくなるにつれて当該ヒントを重要視する度合いが増加し、レベル 1 0 は、当該ヒントを最重要視することを示し、そのヒントに該当する記憶領域の性能は、当該記憶装置各々が提供できる高いものとなる。本実施例では、DMTF の CIM によって定義されるヒントの種類の一例を用いているが、それ以外の CIM で定義されているヒントや CIM 以外のヒントを用いても同様に実施可

能である。

【0019】

本実施例におけるヒントとそれぞれの装置特性情報の対応表を図4(a)、図4(b)に示す。記憶装置301は図4(a)に示される装置特性情報326を、記憶装置302は図4(b)に示される装置特性情報327を、それぞれ、装置特性情報保持部325に保持する。装置特性情報326からは、記憶装置301は、帯域は1Gbpsあるいは2Gbps、ビット単価は1¢/MB、可用性は99.9%である記憶領域を提供可能であることがわかる。装置特性情報327からは、記憶装置302は、帯域は1Gbps、2Gbpsあるいは10Gbps、ビット単価は2¢/MB、可用性は99.999%の記憶領域を提供可能であることがわかる。

【0020】

ビット単価は記憶装置の購入価格と購入容量によって決まってしまうため、本実施例では記憶装置ごとに固有の値としている。また、可用性も装置のハードウェアの構成(電源やコントローラの冗長度や活線挿抜機構の有無)によって決まるため、本実施例では記憶装置ごとに固有の値としている。装置特性情報326や327からも、ヒントのレベルを変えても、割り当てられる記憶領域に変化がないことがわかる。帯域については、FCインタフェースの設定や、帯域制御手段の設定によって可変とすることができるので、記憶装置の中で異なる設定を有する記憶領域を作成することができる。記憶装置301では、AccessBandwidthHintを1~5とすることで帯域が1Gbpsの記憶領域を、AccessBandwidthHintを6~10とすることで帯域が2Gbpsの記憶領域を作成できる。ヒントのレベルが、例えば、ヒントのレベルが1であるものに対し、それよりも大きい6は、記憶領域へのアクセスの帯域がより広いものとしてことができ、アクセスの帯域については、より性能の高い記憶領域を計算機に提供する。記憶装置302では、AccessBandwidthHintを1~5とすることで帯域が1Gbpsの記憶領域を、AccessBandwidthHintを6~8とすることで帯域が2Gbpsの記憶領域を、AccessBandwidthHintを9~10とすることで帯域が10Gbpsの記憶領域を作成できる。記憶装置では、作成した記憶領域を括り付けるFCインタフェースを変えることによって帯域の異なる記憶領域を計算機に提供する。

【0021】

本実施例では、ヒントのレベルである値が異なっても装置特性情報が同じ性能値や性能を示すものである記憶領域を提供する例を含めているが、それぞれのヒントのレベルに対して装置特性情報をそれぞれ異なる値を設定してもよい。

●管理計算機

図12は、実施例1における管理計算機の構成を示す図である。管理計算機600は、プログラムの実行を司るCPU650と、プログラムやプログラムの実行に必要なデータを格納するメモリ660と、プログラムの実行状態を表示するためのディスプレイ680と、管理者の指示を入力するためのキーボード682とマウス684と、第二のネットワークと接続するためのネットワークインタフェース690を設けている。

【0022】

管理計算機600のメモリ660は、記憶装置から装置特性情報と記憶領域情報を取得するための装置情報取得プログラム605と、記憶装置に記憶領域を作成するための記憶領域作成プログラム610と、記憶装置が追加された場合に記憶領域を評価するための記憶領域評価プログラム622と、記憶領域評価プログラム622の評価結果に基づき記憶領域を記憶装置から記憶装置に移動するための記憶領域移動プログラム624とを有している。また、メモリ660には図示していないが、記憶装置から取得する装置特性情報や記憶領域情報、計算機から取得するボリューム管理テーブルも保持する。また、ボリューム移行プログラム630は、実施例2で説明する。

【0023】

なお、プログラムは、管理計算機600内のROMや磁気ディスク等の不揮発性記憶媒体に記憶され、管理計算機の起動時にメモリ360にロードされて実行されるものとする。また

、プログラムの処理を管理計算機内にハードウェア的に構成してもよい。

(2) 記憶装置が実施する記憶領域割当処理の詳細

次に、記憶装置における記憶領域割当プログラム330の処理について説明する。図13に、記憶領域割当プログラム330のフローチャートを示す。記憶領域割当プログラム330は、管理計算機600から記憶領域割当要求を受け取ると(ステップ13010)、記憶領域情報保持部382に保持されている記憶領域情報を更新する(ステップ13020)。具体的には、未割当の記憶領域から要求された容量を切り出し、計算機に割当済みの記憶領域とする(ステップ13025)。さらに、未使用のFCインタフェースから、AccessBandwidthHintの値に対応する未使用のFCインタフェースを検索し、その番号を記憶領域情報に登録する。そして、受け取ったヒントの値を記憶領域情報に登録する。ステップ13025の後、記憶領域情報保持部382に保持している記憶領域情報が更新されたことを通知する(ステップ13030)。この通知先は、記憶領域制御モジュール370と、管理計算機600である。当該通知により、記憶領域制御モジュール370は、記憶領域情報を参照し、作成された記憶領域を割り当てられたFCインタフェースに接続する。また、当該通知により、管理計算機600は、記憶領域割当が完了したことを把握する。

【0024】

図5に記憶領域情報の例を示す。図5(a)では、記憶領域3412がFCインタフェース3714に割当済みで、容量が200GB、帯域が2Gbpsであることを示している。ボリューム割当時に受け取ったヒントは、AccessBandwidthHintが「10」、StorageCostHintが「0」、DataAvailabilityHintが「10」であることがわかる。

(3) 管理計算機600が実施する処理の概要

本実施例において管理計算機600が実施する処理の概要について説明する。本実施例において、管理計算機600は、次のような手順で処理を実施する。

まず、記憶装置301から装置情報の取得する処理を行い、記憶装置301に記憶領域を作成を指示する処理を行う。そして、新しく設置された記憶装置302を検出する処理を実施し、最後に、記憶領域を再評価して移動を指示する処理を行う。

【0025】

上記の各処理を実現するのは、以下の三つのプログラムである。つまり、記憶装置を検出して情報を取得する処理である装置情報取得プログラム605と、記憶装置に記憶領域を作成するよう指示する記憶領域作成プログラム610と、記憶装置に作成した記憶領域を再評価する処理である記憶領域評価プログラム622である。これらのプログラムのフローについて、説明する。

(3-1) 装置情報の取得処理

図14に、管理計算機600が実施する装置情報取得プログラム605のフローチャートを示す。まず、管理計算機600は、第二のネットワーク500に接続されている記憶装置を検出して、装置特性情報を取得する(ステップ14010)。具体的には、第二のネットワーク500に対して一定時間間隔で装置特性情報取得要求を配信し、装置特性情報取得要求に記憶装置が応答して、装置特性情報を返信すると、管理計算機600は、記憶装置を検出したことになる。返信された装置特性情報はメモリ660に保持し、以後、記憶領域の作成処理の際に、どの記憶装置に記憶領域を作成するかの判断に使用する。ステップ14010の後、管理計算機600は、記憶領域情報を取得する(ステップ14020)。具体的には、検出した記憶装置に対して一定時間間隔で記憶領域情報取得要求を発行し、返信された記憶領域情報をメモリ660に保持する。

【0026】

本実施例では、まず記憶装置301しか接続していない場合について説明する。管理計算機600が装置情報取得プログラム605を実行すると、記憶装置301から、装置特性情報326を受信する。

(3-2) 記憶領域の作成処理

●記憶領域作成プログラム610とそのプログラムを適用した具体例その1

図15に、管理計算機600が実施する記憶領域作成プログラム610のフローチャートを示す。図15の処理の具体例を図5、図6及び図8と共に説明する。管理計算機600のCPUがメモリ666に格納されている記憶領域作成プログラム610を実行する。

【0027】

はじめに、管理計算機600は、記憶領域割当要求を作成する(ステップ15010)。具体的には、管理者が、ディスプレイやキーボードを用いて、記憶領域作成対象の記憶装置と、記憶領域を利用する計算機と、記憶領域割当要求パラメータとを入力する。

【0028】

入力画面の例を図8に示す。管理計算機600のディスプレイ680に、図8(a)に示すような記憶領域作成のための設定画面800が表示する。設定画面800には、六つの入力フィールド(810, 812, 818, 820, 822, 824)がある。管理者は、記憶領域へのデータの読みだしや書込みが可能な計算機を計算機入力フィールド810に、希望する記憶領域の容量を容量入力フィールド812に、希望する記憶装置の番号を記憶装置入力818、希望するヒントをヒント820、822、824に、記憶領域割当要求パラメータとして、それぞれの入力フィールドに入力し、作成ボタン830を押す。この操作により、入力された記憶領域割当要求パラメータをもとに記憶領域割り当て要求を作成する。

【0029】

たとえば、「高速・コスト無視・高可用性」を特徴とする記憶領域を作成するために、ここでは、計算機入力フィールド810に「100」、容量入力フィールド812に「200 (GB)」、記憶装置入力フィールド818に「301」、AccessBandwidthHint入力フィールド820に「10」、StorageCostHint入力フィールド822に「0」、DataAvailabilityHint入力フィールド824に「10」を入力し、作成ボタン830を押す。この操作により、入力された記憶領域割当要求パラメータをもとに記憶領域割り当て要求を作成する。図15に戻り、記憶領域作成プログラム610の処理について説明する。

【0030】

次に、管理計算機600は、前記ステップ15010で作成された記憶領域割り当て要求により指定された記憶装置に対して、前記ステップ15010で作成された記憶領域割当要求を送信する(ステップ15020)。

【0031】

ステップ15020の後、記憶装置301では、管理計算機600から記憶領域割当要求を受け、前述した記憶領域割当プログラム330を実行する。具体的には、割り当て要求により200GBの記憶領域3412を切り出し、AccessBandwidthHintが「10」であることから、装置特性情報326を参照し、未使用である2GbpsのFCインタフェース3714に割り当て、記憶領域情報383を図5(a)のように更新する。ここで作成した記憶領域は、帯域は2Gbpsで、ビット単価は1¢/MB、可用性は99.9%である。また、ヒントのレベルが、StorageCostHintが「0」で、AccessBandwidthHintとDataAvailabilityHintが「10」であることは、コストにいとめをつけず(コストを重要視せず)、帯域と可用性が高い(帯域と可用性を重要視した)記憶領域であることを意味している。

【0032】

そして、管理計算機600は、記憶装置から記憶領域割当完了通知を受信したら、記憶領域を作成した記憶装置の更新された記憶領域情報383を取得する(ステップ15030)。

【0033】

最後に、管理計算機600は、前記ステップ15010で指定された計算機に対して、記憶領域作成が完了したのでボリューム管理テーブルを更新するよう要求する(ステップ15040)。例えば、計算機100に対して、ボリューム管理テーブル115を図6(a)に示すように更新するよう指示する。指示を受けた計算機100は、ボリューム管理テーブル115を更新することで、計算機100は、「C」ドライブとして、記憶装置301

の記憶領域 3412 を FC インタフェース 3714 を介して使用することができる。以上が記憶領域作成プログラム 610 である。

【0034】

なお、前記図 8 (a) に示すような記憶領域作成のための設定画面 800 は、図 8 (b) に示すような記憶領域作成のための設定画面 801 であってもよい。設定画面 800 と 801 の違いについて説明する。設定画面 800 では、各管理者は、ある計算機に記憶領域を割り当てるために、記憶装置や 3 つのヒントを含む六つの入力フィールド全てを指定する必要がある。しかしこれは、各管理者に記憶領域作成先の記憶装置やヒントについての高度または膨大な知識を要求するものである。また、複数の記憶装置や複数の計算機の管理を複数の管理者が分担で行っている場合、各管理者が各自の基準で記憶装置やヒントを指定するため、システム全体での記憶領域作成の基準が統一されない恐れがある。設定画面 801 を使う場合は、システム構成を熟知した管理者が、使用頻度の高い記憶領域作成先の記憶装置とヒントの値の組である「ボリュームポリシー」をあらかじめ定義しておき、管理計算機 600 に格納しておく。各管理者は、計算機入力フィールド 810 と、容量入力フィールド 812 と、ボリュームポリシー選択フィールド 816 の三つの入力フィールドを指定する。前記三つの入力フィールドを指定されたとき、管理計算機 600 は、あらかじめ格納されている「ボリュームポリシー」と照合することで、前記設定画面 800 と等しい六つの入力を得ることができる。設定画面 801 により、管理者は、記憶領域作成の入力負担を軽減することができる。

【0035】

また、ボリュームポリシーの例として、「低コスト重視」や「アクセス帯域のみ重視」があげられる。「低コスト重視」のボリュームポリシーの例としては、「StorageCostHint」が 10 であり、他のヒントが、0 と定義されて、ある記憶装置が指定されたボリュームポリシーとして管理計算機 600 のメモリに保持されている。「アクセス帯域のみ重視」のボリュームポリシーの例では、「AccessBandwidthHint」が「10」であり、他のヒントが「0」と定義されており、記憶装置が指定されたボリュームポリシーとして同様にメモリ 660 に保持されている。

●記憶領域作成プログラム 610 適用の具体例その 2

次に、記憶領域作成プログラム 610 の具体例として、記憶装置 301 にさらに別の記憶領域を作成する処理について説明する。

【0036】

同様に、図 8 に示すような記憶領域の作成するための設定画面 800 で、計算機入力フィールド 810 に「100」、記憶装置容量入力フィールド 812 に「200 (GB)」、AccessBandwidthHint 入力フィールド 820 に「1」、StorageCostHint 入力フィールド 822 に「10」、DataAvailabilityHint 入力フィールド 824 に「0」を入力して作成ボタン 830 を押すことにより、管理計算機が、入力されたパラメータにより記憶領域割り当て要求を作成し、記憶領域割り当て要求を記憶装置 301 に送信する。

【0037】

記憶装置 301 は、記憶領域割り当て要求を受け取ると、要求された 200 GB の記憶領域を切り出し、AccessBandwidthHint が「1」であることから、未使用である 1 Gbps の FC インタフェース 3711 を割り当てるように、記憶領域情報 383 を図 5 (b) のように更新する。ここで作成した記憶領域は、帯域は 1 Gbps で、ビット単価は 1 ¢/MB、可用性は 99.9% である。また、StorageCostHint が「10」で、AccessBandwidthHint が「1」であることから、帯域よりもコストを優先していることがわかる。また、DataAvailabilityHint が「0」であることから、可用性についても重要度が低いことがわかる。

【0038】

記憶領域の割り当てが完了すると、管理計算機 600 は、計算機 100 に対して、ボリューム管理テーブル 115 を図 6 (b) に示すように更新するよう指示する。ボリューム管理テーブル 115 を更新することで、計算機 100 は、「D」ドライブとして、記憶装置 301

の記憶領域 3411 を FC インタフェース 3711 を介して使用することができる。

(3-3) 新記憶装置の検出処理

次に、既に作成した二つの記憶領域が記憶装置 301 にある状態で、記憶装置 302 を計算機システムに追加した場合の管理計算機 600 の処理について説明する。本実施例で記憶装置 302 を計算機システムに追加する場合には、FC インタフェース 3721 ~ 3729 を第一のネットワークに、ネットワークインタフェース 390 を第二のネットワークに接続する。

【0039】

管理計算機 600 は、前記装置情報取得プログラム 605 のステップ 14010 を一定時間間隔で実行していることから、記憶装置 302 を計算機システムに追加された段階で、管理計算機 600 のネットワークインタフェース 690 と第二のネットワーク 500 を介して接続する記憶装置 302 を検出し、記憶装置 302 の装置特性情報と、記憶領域情報を取得することが可能である。管理計算機 600 は、取得した装置特性情報をメモリ 660 に保持し、以後、記憶領域の作成処理の際に、どの記憶装置に記憶領域を作成するか判断に使用する。

(3-4) 記憶領域を再評価して移動する処理

管理計算機 600 は、新しい記憶装置を検出した後、すでに割り当て済みの記憶領域の記憶領域評価プログラム 622 を開始する。再評価とは、新しく追加された記憶装置の装置特性情報と、他の記憶装置の記憶領域情報を参照し、よりヒントに合致する記憶領域を提供できる記憶装置がどれか判断することである。提供可能な場合には、実際に新しく追加された記憶装置にヒント情報を用いて記憶領域を作成し、記憶領域に保持されているデータを移動することで、計算機に対して、よりヒントに合致する記憶領域を提供することができる。以下、はじめに管理計算機 600 が実行する記憶領域評価プログラム 622 を説明し、その後本実施例における記憶領域評価プログラム 622 適用の具体例について詳述する。

●記憶領域評価プログラム 622 のフローの説明

図 16 に、管理計算機 600 が実施する記憶領域評価プログラム 622 のフローチャートを示す。

管理計算機 600 は、前記装置情報取得プログラム 605 を用いて最新の装置情報を取得する。

【0040】

取得した全記憶装置の全記憶領域情報のエントリに対して、ステップ 16030 からステップ 16130 までの処理を実施する(ステップ 16020)。

【0041】

まず、未処理の記憶領域情報エントリを選択する(ステップ 16030)。

【0042】

そして、選択された記憶領域情報エントリの全ヒント情報に対して、ステップ 16050 からステップ 16090 までの処理を実施する(ステップ 16040)。

【0043】

まず、未処理のヒント情報を選択する(ステップ 16050)。

【0044】

選択したヒント情報について、現在選択されている記憶領域情報エントリの当該ヒント情報の値をもとに、全記憶装置のヒント情報の値に対応する装置特性情報を取得する(ステップ 16060)。具体的には、現在選択されている記憶領域情報エントリの当該ヒント情報とそのヒント情報の値の組をキーとして、全記憶装置の装置特性情報を検索し、当該ヒント情報の値で提供される装置特性情報を取得すればよい。

【0045】

その後、ステップ 16060 において取得した装置特性情報を比較し、現在選択されている記憶領域情報エントリの属する記憶装置よりも、高い性能を備える記憶領域を提供する記憶装置があるかどうかを判断する(ステップ 16070)。

【0046】

ステップ16070において、もし高い性能を備える記憶領域を提供する記憶装置が存在する場合は、現在記憶領域が存在する記憶装置から、高い性能を備える記憶領域を提供する記憶装置へ、記憶領域を移動させた方が、記憶領域の性能は向上することが期待される。よって、高い性能を備える記憶領域を提供する記憶装置を移動先候補とする(ステップ16080)。ここで、複数の記憶装置が高い性能を備える記憶領域を提供する場合は、最も高い性能を備える記憶領域を提供する記憶装置のみ移動先候補としてもよいし、あるいは、全ての記憶装置を移動先候補としてもよい。また、最も高い性能を備える記憶領域を提供する記憶装置が複数台存在する場合についても、管理計算機600が任意に移動先候補となる記憶装置を1台選択してもよいし、あるいは、全ての記憶装置を移動先候補としてもよい。

【0047】

ステップ16070において、より高い性能を備える記憶領域を提供する記憶装置が存在しない場合は、現在選択中のヒント情報に関しては移動先候補が無いということになる。最後に、選択したヒントを処理済にし(ステップ16090)、ステップ16040に戻る。

前記ステップ16040の繰り返し処理で全ヒント情報に対して処理が実施された後、管理計算機600は、記憶領域移動先の記憶装置を決定する(ステップ16100)。ここで、記憶領域移動先の記憶装置を決定する方法としては、前記ステップ16040の繰り返し処理で移動先候補となった全ての記憶装置から、管理計算機600が任意の記憶装置を選択してもよい。あるいは、移動先候補となった全ての記憶装置の情報をディスプレイ680に出力し、管理者に選択させてもよい。あるいは、あらかじめ管理者が定義して、ヒント情報に重み付けをし、ある特定ヒントで移動先候補となった記憶装置が優先的に選択されるようにしてもよい。また、決定した移動先の記憶装置が、あるヒント情報の観点では性能向上するが、別のヒント情報の観点では性能が向上しないという場合が存在する。例えば、アクセス帯域は向上するが、可用性が低下するとうような事例である。この場合にも、記憶領域を移動してもよいかの確認のためのダイアログを表示して管理者に確認を求めてもよいし、あるいは、管理者に確認を求めずに次ステップに移ってもよい。

【0048】

ステップ16100で記憶領域移動先の記憶装置を決定した後、管理計算機600は、記憶領域を移動する(ステップ624)。

【0049】

ステップ624の後、選択した記憶領域情報エントリを処理済にし(ステップ16130)、ステップ16020に戻る。

【0050】

以上が記憶領域評価プログラム622である。

●記憶領域移動プログラム624のフローの説明

図17に、管理計算機600が、記憶領域評価プログラム622のステップ16100の後において実施する、記憶領域移動プログラム624のフローチャートを示す。

【0051】

まず、管理計算機600は、移動元の記憶領域が、記憶領域受入プログラムを利用しているかどうか判断する(ステップ17010)。記憶領域受入プログラムについては、実施例2で後述するので、本実施例では、ここはNOとなる。

【0052】

次に、移動元の記憶領域が、計算機に登録されているかどうか判断する(ステップ17020)。具体的には、第二のネットワークにボリューム管理テーブル取得要求を配信し、返信のあった全ての計算機のボリューム管理テーブルの全てのボリュームエントリのうち、移動対象の記憶領域の記憶領域番号と一致する記憶領域番号を持つエントリを抽出すればよい。

【0053】

ステップ17020で利用中の計算機があった場合、管理計算機600は、移動対象の記憶領域を登録していた全ての計算機に指示して、当該計算機のボリューム管理テーブルに存在する、移動対象の記憶領域が移動中であると変更するよう指示する(ステップ17030)。具体的には、当該計算機のボリューム管理テーブル115に存在する、移動対象の記憶領域のエントリの、移動中フラグを「0」から「1」に変更する。計算機は、記憶領域をアクセスする度にこのボリューム管理テーブル115の移動中フラグを参照することで、記憶領域の移動処理中のアクセスを抑止することができる。

【0054】

ステップ17010で移動元の記憶領域が記憶領域受入プログラムを利用していた場合、及び、ステップ17020で、移動元の記憶領域が計算機に登録されていない場合、及び、ステップ17030の後に、管理計算機600は、記憶領域移動先となった記憶装置に対して、記憶領域を新規作成する。本ステップは、前記記憶領域作成プログラム610に同じであるため、詳細説明を省略する。

【0055】

記憶領域移動先の記憶装置に記憶領域を新規作成した後、管理計算機600は、記憶領域移動実行手段700に対して、記憶領域のデータ移動を指示する(ステップ17040)。記憶領域とは、一般的に、データを格納する512バイトを単位とするブロックの集合体である。よって、管理計算機600は、記憶領域移動実行手段700に対する指示パラメータとして、移動元の記憶装置と記憶領域番号、移動先の記憶装置と記憶領域番号を記憶領域移動実行手段700に渡すだけでよい。なお、記憶領域移動実行手段700は、受け取ったパラメータを元に、移動元の記憶装置の記憶領域の最初のブロックから最後のブロックまでを、移動先の記憶装置と記憶領域に複写する。本実施例における記憶領域移動実行手段は、例えば、管理計算機600とIPネットワーク500を介して指示パラメータを受け取る独立の計算機上で実行されるプログラムで、プログラムの実行によりファイバチャネル200を介して記憶領域間の複写を行う。

【0056】

そして、管理計算機600は、ステップ17010と同様に、移動対象の記憶領域を利用中の計算機があるかどうかを探し(ステップ17050)、利用中の計算機があった場合は、当該計算機に対して、ボリューム管理テーブルを更新するよう指示する(ステップ17060)。具体的には、移動先の記憶装置の記憶領域情報から、今回移動先記憶領域となった記憶領域のFCインタフェース番号の値と記憶領域番号の値を抽出し、それらの値で移動中フラグが「1」であるドライブレータのエントリの内容を書き換える。そして、当該エントリの移動中フラグを「1」から「0」に戻す。

【0057】

最後に、管理計算機600は、移動元記憶装置の記憶領域を開放する(ステップ17070)。具体的には、移動元記憶装置に対して、記憶領域情報から、移動元記憶領域エントリを削除するよう指示する。

【0058】

以上が記憶領域移動プログラム624である。

●記憶領域評価プログラム622適用の具体例

記憶領域評価プログラム622適用の具体例を説明するため、本実施例において、記憶装置301に記憶領域3411と3412がある状態で、記憶装置302を計算機システムに追加した場合の処理について説明する。記憶装置302を接続したことにより、管理計算機600には、装置特性情報326と327があるものとする。また、記憶装置302には、初期状態として作成済みの記憶領域はないものとする。

【0059】

まず、ステップ605により、管理計算機600は、第二のネットワーク500に接続されている記憶装置それぞれから、装置情報を取得する。本実施例では、記憶装置301からは、図4(a)に示す装置特性情報326と、図5(b)に示す記憶領域情報383が取得する。また、記憶装置302は、図4(b)に示す装置特性情報327と、まだ記憶領域

をひとつも持っていないので、図 7 (a)に示す空の記憶領域情報 384 が取得できる。

【0060】

管理計算機 600 は、次に、ステップ 16020 からステップ 16130 で実施される、記憶領域情報に記載された記憶領域の評価を行う。本実施例では二つの記憶領域情報 383 と 384 があるが、記憶装置 302 から取得した記憶領域情報 384 は空であるので、記憶装置 301 から取得した記憶領域情報 383 の記憶領域 3412 と 3411 を処理する。

【0061】

管理計算機 600 は、まず、記憶領域情報 383 の一行目である記憶領域 3412 の再評価を開始する。はじめに、記憶領域 3412 のヒント情報 AccessBandwidthHint の値を参照する。AccessBandwidthHint の値は「10」である。そして、装置特性情報 326 と装置特性情報 327 の AccessBandwidthHint の値が「10」の時の帯域の値を参照する（ステップ 16060）。すると、記憶装置 301 では帯域が 2 Gbps であるが、記憶装置 302 では帯域が 10 Gbps に高速化されるため、高性能の記憶領域が提供できるので（ステップ 16070）、移動先候補として記憶装置 302 を特定する。そして、選択したヒント情報を処理済にし（ステップ 16090）、次のヒント情報により記憶領域の評価を行う（ステップ 16040）。

【0062】

次に、記憶領域 3412 のヒント情報 StorageCostHint の値を参照する。StorageCostHint の値は「0」であるので再評価対象とはしない。

【0063】

最後に、記憶領域 3412 のヒント情報 DataAvailabilityHint の値を参照する。DataAvailabilityHint の値は「10」である。そして、装置特性情報 326 と装置特性情報 327 の DataAvailabilityHint の値が「10」の時の可用性の値を参照する。すると、記憶装置 301 では可用性が「99.9%」であるが、記憶装置 302 では可用性が「99.999%」であり、記憶装置 302 に記憶領域を移動すれば、可用性を向上できるため、高性能の記憶領域が提供できるので（ステップ 16060）、移動先候補として記憶装置 302 を特定する。そして、選択したヒント情報を処理済にし（ステップ 16090）記憶領域情報の全ヒント情報を参照したので、記憶領域の評価を終了する。

【0064】

以上の評価により、記憶領域 3412 は、記憶装置 301 から記憶装置 302 に移すことにより、帯域と可用性を向上することができることが判断する。

【0065】

よって、記憶領域評価プログラム 622 は、ステップ 16100 に進み、記憶装置 302 を記憶領域移動先の記憶装置と決定し、ステップ 624、つまり、記憶領域移動プログラム 624 で、記憶領域の移動を開始する。記憶領域 3412 の移動にあたって、管理計算機 600 は、記憶装置 301 及び記憶装置 302 の記憶領域情報から、記憶領域 3412 が記憶領域受入プログラムを利用していないと判断する。また、計算機 100 から、ボリューム管理テーブルを取得し、計算機 100 が記憶領域 3412 を利用中であると判断する。そこで、管理計算機 600 は、ボリューム管理テーブル 115 の記憶領域 3412 の行の移動中フラグを「1」に設定するよう指示し、記憶領域 3412 が移動中であるとする（図 6(c)）。

【0066】

次に、管理計算機 600 は、記憶領域 3412 に割り当てられたヒントと容量から、記憶領域割当要求を作成し、記憶装置 302 に発行する。記憶装置 302 の記憶領域割当プログラム 330 は、記憶領域割当要求を受け取ると、要求された 200 GB の記憶領域 3429 を用意し、AccessBandwidthHint が「10」であることから、未使用である 10 Gbps の FC インタフェース 3727 に割り当てるように、記憶領域情報 384 を図 7 (b) のように更新する。ここで作成した記憶領域は、帯域は 10 Gbps で、ビット単価は 2 ¢/MB、可用性は 99.999% である。記憶装置 302 で、記憶領域 3429 の作

成が完了すると、管理計算機600は、記憶領域移動実行手段700に対して、記憶装置301の記憶領域3412内のデータを、記憶装置302の記憶領域3429内に移動するように指示する。記憶領域3429への複写が完了すると、管理計算機600は、計算機100に対して、ボリューム管理テーブル115を図6(d)に示すように更新するよう指示する(移動中フラグも「0」に戻す)。これにより、計算機100の「C」ドライブは、記憶装置302のFCインタフェース3727に接続された記憶領域3429になる。記憶領域3429には、記憶領域移動実行手段700によって、記憶領域3412のデータが移動されているので、計算機100は、記憶領域3412となんら変わらずデータにアクセスすることができる。

【0067】

最後に、管理計算機600は、記憶装置301に対して、記憶領域3412を未割当とするように記憶領域開放要求を発行する。記憶装置301は、記憶領域3412の記憶領域開放要求を受け取ると、記憶領域情報383を図5(c)のように更新する。これにより、記憶領域3412とFCインタフェース3714が開放される。

【0068】

次に、管理計算機600は、ステップ16020に戻り、図6(b)に示されている記憶領域情報383の二行目である記憶領域3411の再評価を開始する(ステップ16030ないし16040)。

【0069】

まず、記憶領域3411のヒント情報AccessBandwidthHintの値を参照する。AccessBandwidthHintの値は「1」である。そして、装置特性情報326と装置特性情報327のAccessBandwidthHintの値が「1」の時の帯域の値を参照する(ステップ16060)。すると、どちらの記憶装置においても1Gbpsであることがわかる(ステップ16070)。このため、記憶装置302は移動先候補とはならない。ヒント情報を処理済とし(ステップ16090)、次のヒントの値を参照する(ステップ16020)。

【0070】

次に、記憶領域3411のヒント情報StorageCostHintの値を参照する。StorageCostHintの値は「10」である。装置特性情報326と装置特性情報327のStorageCostHintの値が「10」の時のビット単価の値を参照する。すると、記憶装置301ではビット単価は2¢/MBで、記憶装置302ではビット単価は1¢/MBであることがわかる。つまり、記憶装置302に移動することで、ビット単価が高くなってしまうため、ここでも、記憶装置302は移動先候補とはならない。最後に、記憶領域3411のヒント情報DataAvailabilityHintの値を参照する。DataAvailabilityHintの値は「0」であるので再評価の対象とはしない。

【0071】

以上の評価により、記憶領域3411は、移動先候補がないため、記憶装置301から移動するメリットがないことが判定される。よって、管理計算機600は、記憶領域3411を移動せずに、ほかに選択する記憶領域情報エントリがないため、記憶領域評価プログラムの処理を終了する。

【0072】

なお、本実施例では、記憶装置302には記憶領域がない状態を初期状態として説明したが、記憶装置302に記憶領域があった場合でも、上記処理にしたがって処理することで、記憶装置302内の記憶領域も再評価の対象となり、装置特性情報とヒント情報により、適切に記憶領域を移動する。

【0073】

また、本実施例では、新しい記憶装置が追加され、装置情報取得プログラム605が新しい記憶装置を検出した場合に再評価を行うように示したが、記憶装置が新たに追加されなくとも、既存の記憶装置の装置情報に変化があった場合に再評価を行ってもよい。たとえば、記憶装置を購入してから時間がたち、計算機の原価償却が進んだ場合には、ビット単価を変更することができる。このとき、記憶装置で自動的に、ないしは、管理者によっ

て、装置特性情報を変更する。装置情報取得プログラム 605 は、その変化を捕らえて記憶領域評価プログラム 622 を実行することにより、StorageCostHint を重要視する記憶領域を、よりビット単価の低い記憶領域に移動することができる。

【0074】

さらに、一定間隔やユーザの入力による指示で、管理計算機 600 が、装置情報取得プログラム 605 を実行し、装置情報を取得してもよい。

【0075】

また、本実施例では、記憶装置が予め装置特性情報をもっていることを前提に説明した。しかし、記憶装置が装置特性情報を持っていない場合には、管理者が管理計算機で当該記憶装置の装置特性情報を作成してもよい。この場合、管理者が装置特性情報を作成（もしくは変更）した場合に、記憶領域評価プログラム 622 を実行することで、本実施例に従って、記憶領域の再評価を行うことができる。

また、本実施例では、記憶装置 301 と 302 から同じフォーマットで表記された図 4 に示すような対応表が取得できるとした。しかし、実際には、記憶装置のベンダが異なるとフォーマットが異なることが考えられる。その場合には、管理計算機に変換プログラムを用意し、対応表のフォーマットを変換することで、本実施例記載の処理が可能となる。

また、実施例での記憶領域評価プログラムのステップ 16020 では、管理計算機とネットワークを介して接続している全ての記憶装置の記憶領域情報について本プログラムを実行するとしているが、少なくとも一つの記憶装置が保持する記憶領域情報のヒント情報に対して処理を行ってもよい。

【実施例 2】

【0076】

(1) システム構成

図 9 に実施例 2 のシステム構成を示す。実施例 2 では、記憶装置 301 を有する計算機システムに、記憶装置 303 を追加する場合を説明する。

【0077】

本実施例と実施例 1 との違いについて説明する。

第一の違いは、本実施例では、記憶装置 303 を追加する場合は、記憶装置 301 を一度第一のネットワーク 200 から切り離し、記憶装置 303 に接続することである。

第二の違いは、実施例 1 では、記憶領域移動実行手段 700 を第一のネットワーク 200 に接続していたのに対して、本実施例では、記憶領域移動実行手段 700 を記憶装置 303 内に設けていることである。

第三の違いは、図 12 に示す管理計算機 600 の構成である。本実施例では、実施例 1 の構成に加えて、ボリューム移行プログラム 630 をさらに有している。

第四の違いは、本実施例における、追加する記憶装置 303 の構成である。以下で詳しく説明する。

【0078】

記憶装置 303 は、記憶装置 301 を接続するための FC インタフェース 3731～3736 を有する。なお、FC インタフェース 3731～3733 は 1 Gbps で、FC インタフェース 3734～3736 は 2 Gbps とする。

【0079】

さらに、記憶装置 303 は、FC インタフェース 3731～3736 から記憶装置 303 が認識可能な記憶装置 301 の記憶領域を、あたかも記憶装置 303 の記憶領域として計算機 100 に提供することができる機能を有する。このような機能は、記憶装置 303 の記憶領域制御モジュール 380 が、図 10 に示すような記憶領域情報 385 を有することにより可能である。図 10 に示す記憶領域情報 385 は、実施例 1 の記憶領域情報に加えて、記憶領域が格納されている外部記憶装置（本実施例では記憶装置 301）の FC インタフェース番号と記憶領域番号を記憶領域情報と対応付けて保持できるようにしている。

さらに、記憶装置 303 は、計算機 100 と接続する第一のネットワーク 200 と接続

するために、3つのFCインタフェース3737～3739を有する。FCインタフェース3737～3739は、10GbpsのFCインタフェースである。

さらに、記憶装置303は、メモリ360に、記憶領域受入プログラム340を有する。

以上が、本実施例2と実施例1との違いである。

【0080】

なお、実施例2において、図9に示す装置特性情報保持部に保持される装置特性情報は、記憶装置301の装置特性情報が図4(a)、記憶装置303の装置特性情報が図4(b)、という対応関係にあるものとする。また、記憶装置303が有する他の構成のうち、実施例2での言及しないものは、実施例1の図2で既に説明したものと同様であるとする。

(2) 記憶装置が実施する記憶領域割当プログラム

本実施例で記憶装置が実施する記憶領域割当プログラムは、実施例1と変更がない。よって説明を省略する。

(3) 管理計算機600が実施する処理の詳細

本実施例において管理計算機600が実施する処理の詳細について説明する。本実施例において、管理計算機600は、次のような手順で処理を実施する。まず記憶装置301から装置情報の取得する処理を行い、記憶装置301に記憶領域を作成する処理を行う。そして、新しく設置された記憶装置303と記憶装置301を接続する処理を行う。さらに、記憶装置303を検出する処理を実施し、最後に、記憶領域を再評価して移動する処理を行う。

(3-1) 装置情報の取得処理

本実施例で、管理計算機600が、記憶装置の情報を取得するために用いる、装置情報取得プログラム605は、実施例1に同じである。よって説明を省略する。

【0081】

本実施例でも、実施例1と同様に、まずはじめは記憶装置301しか接続していない状態を考える。よって、管理計算機600が装置情報取得プログラム605を実行すると、記憶装置301から、装置特性情報326が送信される。

(3-2) 記憶領域の作成処理

本実施例で、管理計算機600が、記憶領域を作成するために用いる、記憶領域作成プログラム610は、実施例1に同じである。よって説明を省略する。

【0082】

本実施例でも、実施例1と同様に、管理計算機600により、記憶装置301に対して、記憶領域3411と記憶領域3412が作成される。

(3-3) 記憶装置301と記憶装置303の接続処理

先にも述べたように、本実施例では、記憶装置303を追加する場合は、記憶装置301と記憶装置303の接続処理が必要になる。この接続処理は、以下の手順で実施される。

【0083】

まず、記憶装置301と記憶装置303のインタフェースの接続を変更し、記憶装置303から、記憶装置301の装置情報や記憶領域が取得できるようにする。

【0084】

次に、記憶装置303に設けた記憶領域受入プログラム340を実行し、記憶装置301の記憶領域を、あたかも記憶装置303の記憶領域として計算機100が利用できるようにする。

【0085】

最後に、管理計算機600に設けたボリューム移行プログラム630を実行し、計算機100が、記憶装置303の記憶領域を利用するように構成変更する。

【0086】

以下、各手順について詳細を説明する。

●記憶装置のインタフェースの接続変更

本実施例では、記憶装置301を一度第一のネットワーク200から切り離し、記憶装置303に接続する。記憶装置301のFCインタフェース3711～3713は、記憶装置303のFCインタフェース3731～3733に接続される。記憶装置301のFCインタフェース3714～3716は、記憶装置303のFCインタフェース3734～3736に接続される。すなわち、帯域が1GbpsのFCインタフェースどうし、帯域が2GbpsのFCインタフェースどうしのように、同じ帯域どうしのFCインタフェースを互いに接続する。なお、異なる帯域を持つFCインタフェースを接続してもよい、その場合は、帯域の値が小さい方の帯域で動作が行われる。さらに、FCインタフェースどうしを直接接続するのではなく、第一のネットワーク200を介して互いに接続されてもよい。

【0087】

さらに、記憶装置303のネットワークインタフェース390は、第二のネットワーク500に接続される。

●記憶領域受入プログラム340の実行

まず、記憶領域受入プログラム340のフローについて説明する。

【0088】

図18に、記憶装置303が実施する記憶領域受入プログラム340のフローを示す。

【0089】

まず、記憶装置303は、一定間隔でFCインタフェース3731～3736を監視し、新たな記憶装置が接続されていないかをチェックする(ステップ18010)。

【0090】

ステップ18010において、新たな記憶装置が接続されたことを検出したとき、記憶装置303は、第二のネットワーク500を介して、接続された記憶装置に対して、記憶領域情報を要求し、当該装置の記憶領域情報を取得する(ステップ18020)。

【0091】

最後に、記憶装置303は、前記ステップ18020で取得した記憶領域情報をもとに、自記憶装置、つまり、記憶装置303の記憶領域情報を更新する(ステップ18030)。具体的には、以下で後述する。なお、この更新処理を行うにあたり、記憶装置303は、管理者に確認を求めてもよいが、本実施例では、管理者に確認を求めないものとする。

【0092】

以上が記憶領域受入プログラム340である。

【0093】

次に、本記憶領域受入プログラム340の本実施例での適用について説明する。

【0094】

記憶装置のインタフェースの接続変更が完了したのち、記憶装置303は、ステップ18010により、一定間隔にFCインタフェース3731～3736を監視し、記憶装置が接続されていないかをチェックする。

【0095】

記憶装置301が接続されたことを確認すると、記憶装置303は、ステップ18020により、記憶装置301から記憶領域情報383を取得する。

【0096】

記憶装置301の記憶領域情報を取得した記憶装置303は、最後に、ステップ18030により、自装置の記憶領域情報385を更新する。具体的には、記憶装置301の記憶領域3411と3412を、あたかも記憶装置303の記憶領域として計算機100に提供することができるよう、記憶領域情報385に登録する。ステップ18030によって更新された記憶領域情報385を図10(a)に示す。図10(a)のうち、記憶装置303が提供する記憶領域3431が、FCインタフェース3737に接続されており、容量は200GBを示しているのは、図5(a)と同様である。図10(a)では、さらに、記憶装置303の記憶領域の実体が、記憶装置303と接続された外部記憶装置にある場合に対応

するため、外部記憶装置の情報を保持する。例えば記憶領域3431の場合、記憶領域の実体としては、FCインタフェース番号3711に接続された外部記憶装置である記憶装置301の記憶領域番号3411の記憶領域を利用していることがわかる。また、同様に、記憶装置303が提供する記憶領域3432は、FCインタフェース3738に接続されており、容量は200GBであり、さらに、記憶領域の実体としては、FCインタフェース番号3714に接続された外部記憶装置である記憶装置301の記憶領域番号3412の記憶領域を利用していることがわかる。すなわち、記憶装置303が有する記憶領域情報385は、自装置における記憶領域番号、当該記憶領域の容量、当該記憶領域に接続するFCインターフェース番号とその帯域、当該記憶領域のヒント情報と、対応する外部記憶装置の記憶領域の記憶領域番号と外部記憶装置の記憶領域に接続しているFCインターフェース番号を有している。

● ボリューム移行プログラム630の実行

まず、ボリューム移行プログラム630のフローについて説明する。

【0097】

図19に、管理計算機600が実施するボリューム移行プログラム630のフローを示す。

まず、管理計算機600は、装置情報取得プログラム605を実行し、記憶装置の記憶領域情報を取得する(ステップ605)。

【0098】

次に、第二のネットワーク500に計算機のボリューム管理テーブル取得要求を配信し、各計算機からボリューム管理テーブルを取得する(ステップ19020)。

【0099】

最後に、前記ステップ605とステップ19020で取得した、記憶装置の記憶領域情報と、計算機のボリューム管理テーブルから、計算機にボリューム管理テーブルの更新を指示する(ステップ19030)。本ステップの詳細は具体例で説明する。

【0100】

以上がボリューム移行プログラム630である。

【0101】

次に、本ボリューム移行プログラム630の本実施例での適用について説明する。

【0102】

まず、管理計算機600は、ステップ605で、記憶装置301と記憶装置303から記憶領域情報を取得する。具体的には、記憶装置301から、図5(b)に示す記憶領域情報383を取得し、記憶装置303から、図10(a)に示す記憶領域情報385を取得する。

【0103】

次に、管理計算機600は、ステップ19020で、計算機100からボリューム管理テーブルを取得する。具体的には、計算機100から図6(b)に示すボリューム管理テーブル115を取得する。

【0104】

最後に、管理計算機600は、ステップ19030で、計算機100に対して、ボリューム管理テーブル115の更新を指示する。具体的には、まず、記憶装置301の記憶領域情報とボリューム管理テーブル115に含まれるボリューム管理情報から、計算機100が、記憶装置301の記憶領域3411と記憶領域3412を利用していたことを把握する。次に、記憶装置301と記憶装置303それぞれの記憶領域情報から、記憶装置301の記憶領域3411の内容が、記憶装置303から記憶領域3431として提供されていること、また、記憶装置301の記憶領域3412の内容が、記憶装置303から3432として提供されていることを把握する。そこで、管理計算機600は、計算機100に対して、ボリューム管理テーブル115を、図6(b)から図11のように変更するように指示する。変更した図11に示すボリューム管理テーブル115を用いて、計算機100は、今までどおり、記憶装置301の記憶領域3411と記憶領域3412の内容を、

あたかも、記憶装置 303 の記憶領域 3431 と記憶領域 3432 であるかのように利用することができる。

(3-4) 新記憶装置の検出処理

本実施例で、管理計算機 600 が、記憶装置 303 を検出するために用いる、装置情報取得プログラム 605 は、実施例 1 に同じである。よって説明を省略する。

【0105】

本実施例でも、管理計算機 600 は、記憶装置 303 を検出し、装置特性情報 327 を取得する。

(3-5) 記憶領域を再評価して移動する処理

●記憶領域の再評価プログラム及び移動プログラムのフロー

本実施例で、管理計算機 600 が、記憶領域を再評価するために用いる記憶領域評価プログラム 622 は、実施例 1 に同じである。よって、プログラムフローの説明は省略する。

【0106】

本実施例で、管理計算機 600 が、記憶領域を移動するために用いる記憶領域移動プログラム 624 は、実施例 1 に同じである。ただし、ステップ 17010 において、移動元の記憶領域が記憶装置 303 の記憶領域受入プログラムを利用していると判断される可能性がある。本ステップ 17010 での判断方法は、記憶装置 303 の記憶領域情報の外部記憶装置フィールドに、移動元の記憶領域が登録されている場合は、記憶領域受入プログラムを利用していると判断する。また、記憶領域受入プログラムを利用している移動元記憶領域が計算機に登録されているかどうかを判断する場合は、移動元記憶領域の記憶領域番号と FC インタフェース番号を、移動元記憶装置のものではなく、記憶領域受入プログラムを有する記憶装置の記憶領域番号と FC インタフェース番号に読み替える。

【0107】

以降では、記憶装置 301 の記憶領域 3412 と記憶領域 3411 の再評価及び移動の具体例について説明する。

●本実施例での記憶領域の再評価と移動の具体例その 1:

管理計算機 600 は、記憶領域情報 385 の一行目である記憶領域 3412 の再評価を開始する (ステップ 16020)。

【0108】

まず、記憶領域 3412 のヒント情報 AccessBandwidthHint の値を参照する。AccessBandwidthHint の値は「10」である。

【0109】

次に、記憶装置 301 及び記憶装置 303 からそれぞれ取得した装置特性情報 326 と 327 における、AccessBandwidthHint の値が「10」を参照すると、記憶装置 301 では帯域が 2 Gbps であるが、記憶装置 303 では帯域 10 Gbps に高速化されることがわかる (ステップ 16060)。

【0110】

よって、記憶装置 301 の記憶領域 3412 は、外部記憶装置として記憶装置 303 を介してホストに記憶領域を提供するよりも、記憶領域自体を記憶装置 303 に移動することによって、帯域 10 Gbps をフルに生かした記憶領域を提供できるので (ステップ 16060)、移動先候補として記憶装置 303 を特定する。そして、選択したヒント情報を処理済にし (ステップ 16090)、次のヒント情報により記憶領域の評価を行う (ステップ 16040)。

【0111】

次に、管理計算機 600 は、記憶領域 3412 のヒント情報 StorageCostHint の値を参照する。StorageCostHint の値は「0」であるので、このヒント情報については、再評価対象とはしない。

【0112】

最後に、管理計算機 600 は、記憶領域 3412 のヒント情報 DataAvailabilityHint の

値を参照する。DataAvailabilityHintの値は「10」である。

【0113】

次に、記憶装置301及び記憶装置303からそれぞれ取得した装置特性情報326と327における、DataAvailabilityHintの値が「10」のときの可用性の値を参照すると、記憶装置301では可用性が「99.9%」であるが、記憶装置303では可用性が「99.999%」となり（ステップ16060）、記憶領域自体を記憶装置303に移動することによって、可用性を向上でき、高性能な記憶領域を提供できるので（ステップ16070）、移動先候補として記憶装置303を特定する。そして、選択したヒント情報を処理済にし（ステップ16090）記憶領域情報の全ヒント情報を参照したので、記憶領域の評価を終了する。

【0114】

以上の評価により、記憶領域3412は、記憶装置301から記憶装置303に移すことにより、帯域と可用性を向上することができると判断できる。

【0115】

よって、管理計算機600は、データの移動を開始する。本実施例では、評価から移動まで、管理者に確認を求めず実行するが、実施例1と同様に、管理者に確認を求めてもよい。

【0116】

記憶装置301から記憶装置303へ記憶領域を移動することになった、記憶領域3412について、管理計算機600は、記憶領域移動プログラム624を実行する。

【0117】

記憶領域3412は、記憶領域受入プログラムを利用している記憶領域であるので、ステップ17030のような、計算機のボリューム管理テーブルの更新は発生しない。よって、計算機100のボリューム管理テーブルは図11のまま変更されない。

【0118】

次に、管理計算機600は、記憶領域3412を移動するために、記憶領域3412に割り当てられたヒントと容量から、記憶領域割当要求を作成し、記憶装置303に発行する。

【0119】

記憶装置303は、記憶領域割当プログラム330を実行し、要求された200GBの記憶領域3439を用意する。この時点では、記憶領域3439には、FCインタフェースを割り当てないため、FCインタフェース番号には「0」を与えておく（図10(b)）。記憶領域3439の作成が完了すると、記憶装置303から管理計算機へ完了通知を送る。

【0120】

管理計算機600は、記憶装置303からの記憶領域作成完了通知を受け、記憶装置303の記憶領域移動実行手段700に対して、記憶領域3412内のデータを、記憶領域3439内に移動するように指示する。記憶領域移動実行手段700は、記憶領域3412の最初のブロックから最後のブロックまでを、記憶領域3439に複写する。本実施例における記憶領域移動実行手段は、記憶装置上で実行されるプログラムで、プログラムの実行によりファイバチャネル200を介して記憶領域間の複写を行う。

【0121】

記憶領域3439への複写が完了すると、管理計算機600は、記憶装置303の記憶領域情報385を図10(c)に示すように更新する。ここで、記憶装置303は、新たに作成した記憶領域3439をFCインタフェース3738に割り当て、記憶領域番号を「3432」に付け替える。こうすることにより、実施例1で必要であった、移動時の計算機100のボリューム管理テーブル115におけるFCインタフェース番号と記憶領域番号の書き換えは、実施例2では不要になる。

【0122】

最後に、管理計算機600は、記憶領域開放プログラム612を実行し、記憶装置30

1に対して、記憶領域3412を未割当とする。

【0123】

以上で、記憶領域3412の再評価及びその評価に基づく領域移動処理が終了する。

●本実施例での記憶領域の再評価と移動の具体例その2

次に、管理計算機600は、記憶領域情報385の二行目である記憶領域3411の再評価を開始する。

【0124】

はじめに、記憶領域3411のヒント情報AccessBandwidthHintの値を参照する。AccessBandwidthHintの値は「1」である。そして、記憶装置301及び記憶装置303からそれぞれ取得した装置特性情報326と装置特性情報327のAccessBandwidthHintの値が「10」のときの帯域の値を参照する。どちらの記憶装置においても1Gbpsであることがわかる。このため、移動する必要はない。

【0125】

次に、記憶領域3411のヒント情報StorageCostHintの値を参照する。StorageCostHintの値は「10」である。そして、記憶装置301及び記憶装置303からそれぞれ取得した装置特性情報326と装置特性情報327のStorageCostHintの値が「10」のときのビット単価を参照する。記憶装置301ではビット単価は2¢/MBで、記憶装置303ではビット単価は1¢/MBであることがわかる。よって、記憶装置303に移動することで、ビット単価が高くなってしまうため、記憶領域を移動するメリットがないことがわかる。

【0126】

最後に、記憶領域3411のヒント情報DataAvailabilityHintの値を参照する。DataAvailabilityHintの値は「0」であるので、再評価の対象とはしない。

【0127】

以上の評価により、記憶領域3411は、記憶装置301から記憶装置303に移動するメリットがないことが判定される。

【0128】

よって、管理計算機600は、記憶領域3411を移動する記憶領域移動プログラム624を実行しない。

【0129】

実施例2においても、実施例1と同様に、管理計算機600を設けて説明したが、実施例2で設けていた管理計算機600の装置情報取得プログラム605、記憶領域作成プログラム610、記憶領域開放プログラム612、記憶領域情報取得プログラム620、記憶領域評価プログラム622、記憶領域移動プログラム624を記憶装置303に設けてもよい。これにより記憶装置303が、FCインタフェース3731～3736を定期的にチェックし、記憶装置が接続されたことを検知すると、再評価のための一連のシーケンスを実行するように構成することができる。

【0130】

また、管理計算機と記憶装置303間のネットワークの負荷を削減することができる。

【0131】

最後に、実施例1及び実施例2で用いられる領域移動実行手段700について説明する。

【0132】

図20は、図1中の領域移動実行手段700のハードウェア構成図を示す図である。記憶領域実行手段700は、プログラムの実行を司るCPU710と、プログラムやプログラムの実行に必要なデータを格納するメモリ720と、第二のネットワークと接続され、管理計算機600とのデータの授受を行うネットワークインタフェース730と、記憶装置とのデータの授受を行うFCインタフェース750を設けている。メモリ720には、記憶領域移動手段700が管理計算機600の指示により実行する記憶領域複写プログラム740が格納されている。

【0133】

図21は、図1中の領域移動実行手段700が実施する記憶領域複写プログラム740のフローを示す図である。領域移動実行手段700のCPU710が、メモリ720に格納されている記憶領域複写プログラム740を実行する。

【0134】

まず、領域移動実行手段700は、管理計算機600から、記憶領域を複写するためのパラメータとして、移動元の記憶装置と記憶領域番号、移動先の記憶装置と記憶領域番号を受け取る(ステップ21010)。

【0135】

ステップ21010で受け取ったパラメータをもとに、領域移動実行手段700のCPU710は、第一のネットワークを介して移動元の記憶装置の記憶領域の最初のブロックから最後のブロックまでを、移動先の記憶装置と記憶領域に複写する(ステップ21020)。

【0136】

具体的には、まず、CPU710が、FCインタフェース750を介して、ステップ21010のパラメータに含まれる移動元の記憶装置の記憶領域に対してSCSIのReadコマンドを送信し、該当する記憶領域のあるブロックのデータを読み取る。そして、CPU710は、FCインタフェース750を介して、パラメータに含まれる移動先の記憶装置の記憶領域に対して、対応するブロックへのSCSIのWriteコマンドを、読み取ったデータとともに送信する。Writeコマンドを受け取った移動先の記憶装置は、データを該当する記憶領域のブロックに書き込む。このReadコマンドとWriteコマンドの処理を、記憶領域の最初のブロックから最後のブロックまで繰り返せばよい。なお、領域移動実行手段700は、読み取ったデータをメモリ720に格納し、メモリから読み出して、移動先の記憶装置にWriteコマンドとともに送信してもよい。

【0137】

さらに、図21には図示しないが、CPU710は、ステップ21020を実行するにあたり、移動元の記憶装置と記憶領域と、移動先の記憶装置と記憶領域が、FCインタフェース750からアクセス可能であることを確認するものとする。もしアクセス可能でない場合は、記憶領域複写プログラム740を中断し、管理計算機600にアクセス可能でないことを通知する。通知を受けた管理計算機600は、管理者に対して、アクセス可能な設定をするよう警告をディスプレイ680に表示し、設定変更を促す。設定変更実施後、キーボード682やマウス684などの入力による管理者からの指示に応じて、管理計算機600から領域移動実行手段700へ記憶領域複写プログラム740を再開を指示する。

【0138】

なお、実施例1の図2の計算機システムにおいて、領域移動実行手段700は、記憶装置301、記憶装置302、計算機100、及び管理計算機600とは独立した異なるハードウェアとしている。それにより、各装置、特に、管理計算機600の処理の負荷を削減することができる。

【0139】

しかし、記憶装置301、記憶装置302、計算機100、または管理計算機600のいずれかの装置が、記憶領域複写プログラム740を具備して記憶領域複写プログラム740を実行することにより、図20に示すような外部装置としての領域移動手段700を含まない計算機システムの構成であってもよい。ただし、管理計算機600のCPU650が記憶領域複写プログラム740を実行する場合、図9には図示していないが、管理計算機600が、第1のネットワークを介して記憶装置とのデータの授受を行うFCインタフェースを備え、CPU650が図21で示した処理を行う。

【0140】

また、実施例2の図9の計算機システムにおいて、領域移動実行手段700は、記憶装置303内に具備されている。しかし、記憶装置301、計算機100、及び管理計算機600のいずれかの装置が、ハードウェア的あるいはソフトウェア的に領域移動手段70

0を具備してもよいし、領域移動実行手段700が独立したハードウェア構成を有する外部装置であってもよい。

【0141】

以上説明した種々の実施例では、ネットワークを介する複数の記憶装置が割り当てた膨大なボリュームを有する大規模なストレージシステムの管理者にとって、リソース管理負荷の低減例えば、ボリューム（記憶領域）の移行の判断などの管理負荷の低減が実現される。

【0142】

また、記憶装置ごとに決められる記憶装置の単位時間当たりの使用時間（使用率）の上限値や記憶領域に対するアクセス種別のほかに、記憶領域の信頼性やコストなど実測では取りえない情報、すなわち、実測では得ることのできない予め記憶領域に割り当てられた「ヒント」に基づいても記憶領域を再評価することができる。

本明細書で説明した実施例により、SANなどの複数の記憶装置を有する計算機システムにおいて、記憶装置の追加やSANに属する記憶装置の構成変更を検出して、既存の記憶領域を、記憶領域作成の際に与えられる「ヒント」に基づいて再配置する方法を提供し、記憶領域の配置を管理し、リソースの有効な利用をすることができる。

【図面の簡単な説明】

【0143】

【図1】本発明を実施するための形態の概要を示す図である。

【図2】実施例1における計算機システムを示す図である。

【図3】記憶領域割当要求のフォーマットを示す図である。

【図4】装置特性情報を示す例である。

【図5】記憶装置301における記憶領域情報の例を示す図である。

【図6】計算機のボリューム管理テーブルを示す図である。

【図7】記憶装置302における記憶領域情報を示す図である。

【図8】記憶領域作成のための設定画面を示す図である。

【図9】実施例2における計算機システムを示す図である。

【図10】記憶装置303における記憶領域情報を示す図である。

【図11】実施例2における計算機のボリューム管理テーブルを示す図である。

【図12】実施例における管理計算機の構成を示す図である。

【図13】実施例において記憶装置が実施する記憶領域割当プログラムのフローチャートである。

【図14】実施例において管理計算機が実施する装置情報取得プログラムのフローチャートである。

【図15】実施例において管理計算機が実施する記憶領域作成プログラムのフローチャートである。

【図16】実施例において管理計算機が実施する記憶領域評価プログラムのフローチャートである。

【図17】実施例において管理計算機が実施する記憶領域移動プログラムのフローチャートである。

【図18】実施例2において記憶装置303が実施する記憶領域受入プログラムのフローチャートである。

【図19】実施例2において管理計算機が実施するボリューム移行プログラムのフローチャートである。

【図20】実施例における記憶領域移動手段のハードウェア構成を示す図である。

【図21】実施例において記憶領域移動手段が実施する記憶領域複写プログラムのフローチャートである。

【符号の説明】

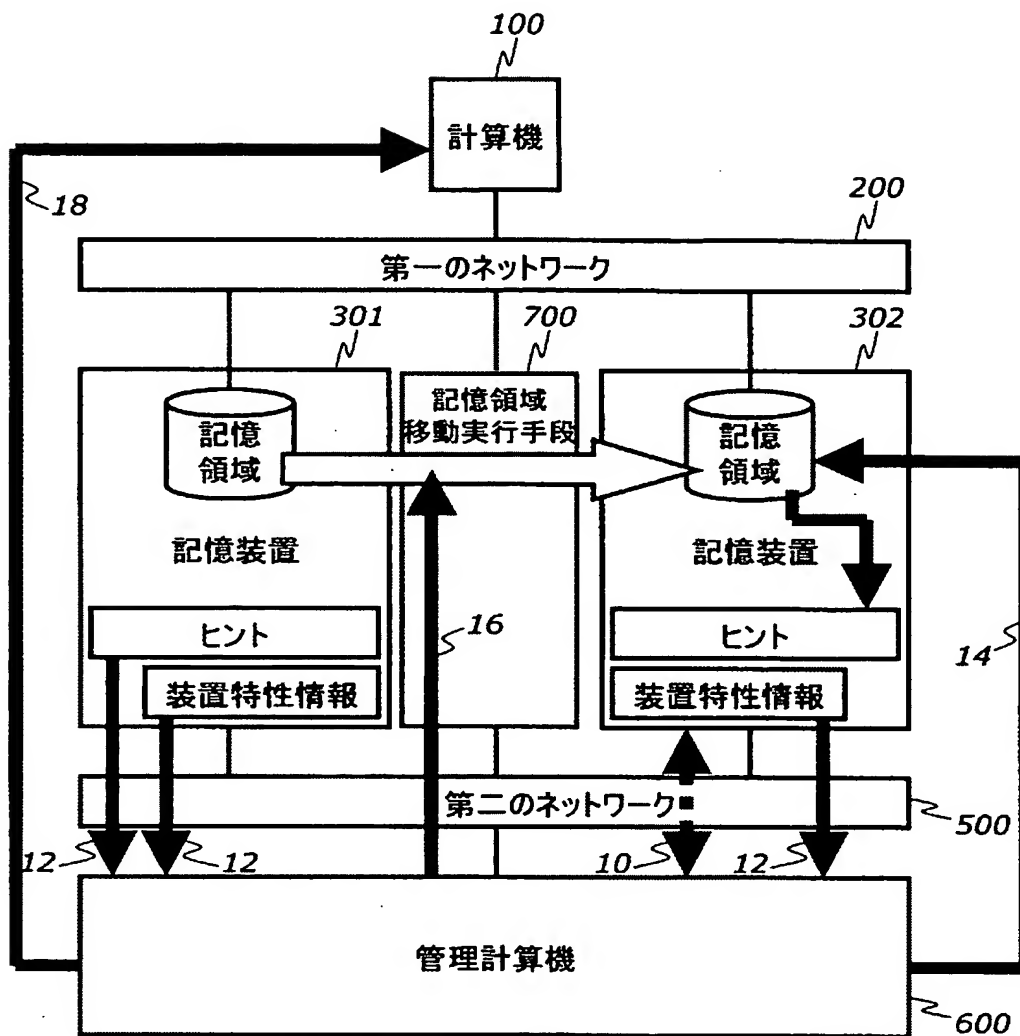
【0144】

100…計算機, 105…CPU, 110…メモリ110, 115…ボリューム管理テー

ブル, 1 2 0…ファイバチャネルインタフェース, 1 3 0…ネットワークインタフェース, 2 0 0…第一のネットワーク, 3 0 1…記憶装置, 3 0 2…記憶装置, 3 0 3…記憶装置, 3 2 5…装置特性情報保持部, 3 2 6…装置特性情報, 3 2 7…装置特性情報, 3 3 0…記憶領域割当プログラム, 3 4 0…記憶領域受入プログラム, 3 5 0…CPU, 3 6 0…メモリ, 3 8 0…記憶領域制御モジュール, 3 8 2…記憶領域情報保持領域, 3 8 3…記憶領域情報 (記憶装置 3 0 1 用), 3 8 4…記憶領域情報 (記憶装置 3 0 2 用), 3 8 5…記憶領域情報 (記憶装置 3 0 3 用), 3 8 9…キャッシュ, 3 9 0…ネットワークインタフェース, 5 0 0…第二のネットワーク, 6 0 0…管理計算機, 6 0 5…装置情報取得プログラム, 6 1 0…記憶領域作成プログラム, 6 2 2…記憶領域評価プログラム, 6 2 4…記憶領域移動プログラム, 6 3 0…ボリューム移行プログラム, 6 5 0…CPU, 6 6 0…メモリ, 6 8 0…ディスプレイ, 6 8 2…キーボード, 6 8 4…マウス, 6 9 0…ネットワークインタフェース, 7 0 0…記憶領域移動実行手段, 3 4 1 1…記憶領域, 3 4 1 2…記憶領域, 3 4 2 9…記憶領域, 3 4 3 1…記憶領域, 3 4 3 2…記憶領域, 3 4 3 9…記憶領域, 3 7 1 1～3 7 1 6…FCインタフェース, 3 7 2 1～3 7 2 9…FCインタフェース, 3 7 3 1～3 7 3 9…FCインタフェース

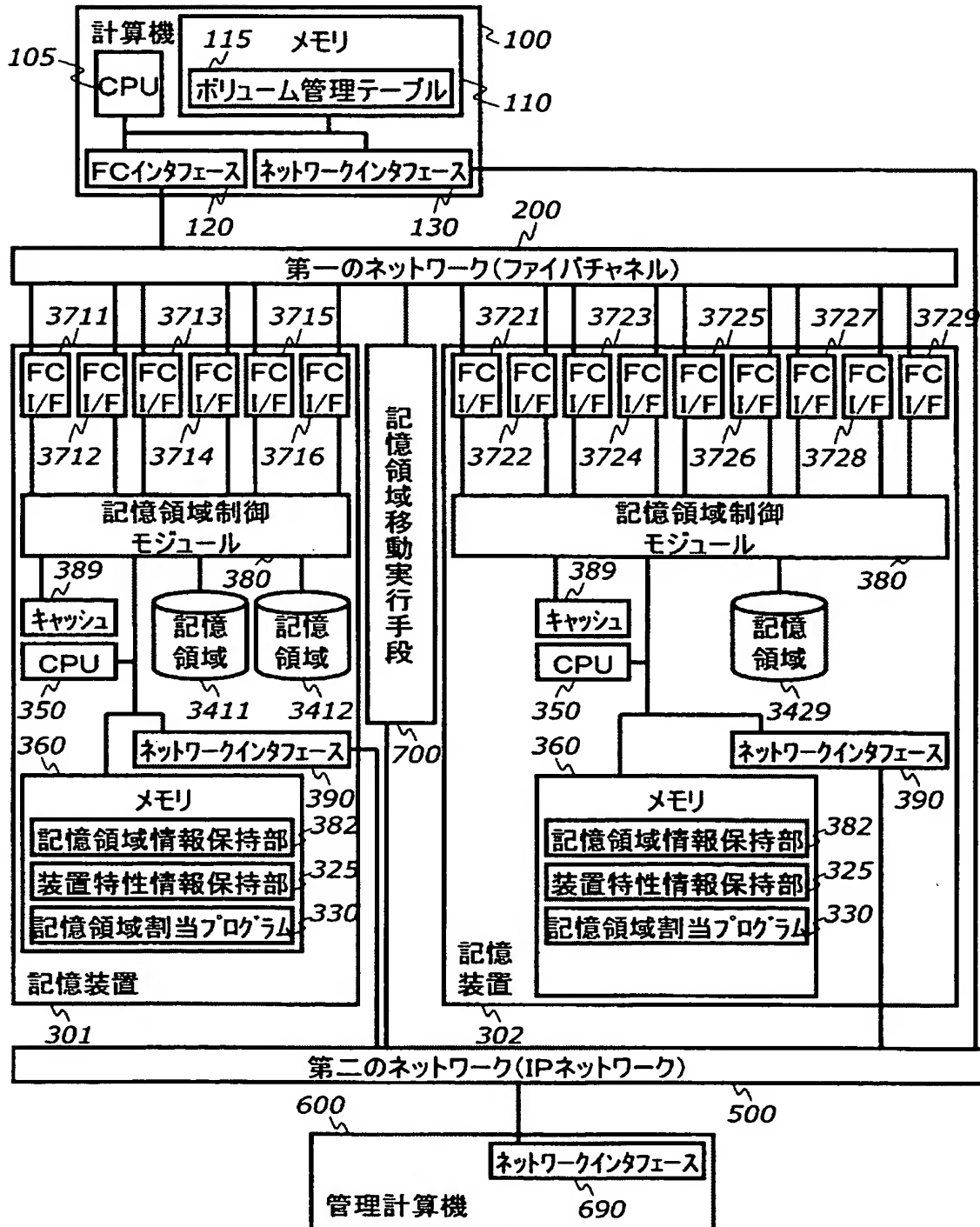
【書類名】 図面
【図 1】

図 1



【図 2】

図 2



【図 3】

図 3

記憶領域割当要求	
容量	
AccessBandwidthHint	
StorageCostHint	
DataAvailabilityHint	

【図 4】

図 4

図 4 (a)

326
N

Hint	Access Bandwidth	Storage Cost	Data Availability
0	Don't Care	Don't Care	Don't Care
1	1Gbps	1 ¢ /MB	99.9%
2	1Gbps	1 ¢ /MB	99.9%
3	1Gbps	1 ¢ /MB	99.9%
4	1Gbps	1 ¢ /MB	99.9%
5	1Gbps	1 ¢ /MB	99.9%
6	2Gbps	1 ¢ /MB	99.9%
7	2Gbps	1 ¢ /MB	99.9%
8	2Gbps	1 ¢ /MB	99.9%
9	2Gbps	1 ¢ /MB	99.9%
10	2Gbps	1 ¢ /MB	99.9%

図 4 (b)

327
N

Hint	Access Bandwidth	Storage Cost	Data Availability
0	Don't Care	Don't Care	Don't Care
1	1Gbps	2 ¢ /MB	99.999%
2	1Gbps	2 ¢ /MB	99.999%
3	1Gbps	2 ¢ /MB	99.999%
4	1Gbps	2 ¢ /MB	99.999%
5	1Gbps	2 ¢ /MB	99.999%
6	2Gbps	2 ¢ /MB	99.999%
7	2Gbps	2 ¢ /MB	99.999%
8	2Gbps	2 ¢ /MB	99.999%
9	10Gbps	2 ¢ /MB	99.999%
10	10Gbps	2 ¢ /MB	99.999%

【図5】

図5

ABH: AccessBandwidthHint

SCH: StorageCostHint

DAH: DataAvailabilityHint

図5 (a)

383
N

記憶領域番号	FC I/F 番号	容量	帯域	記憶領域のヒント情報		
				ABH	SCH	DAH
3412	3714	200GB	2Gbps	10	0	10

図5 (b)

383
N

記憶領域番号	FC I/F 番号	容量	帯域	記憶領域のヒント情報		
				ABH	SCH	DAH
3412	3714	200GB	2Gbps	10	0	10
3411	3711	200GB	1Gbps	1	10	0

図5 (c)

383
N

記憶領域番号	FC I/F 番号	容量	帯域	記憶領域のヒント情報		
				ABH	SCH	DAH
3411	3711	200GB	1Gbps	1	10	0

【図6】

図6

115
N

図6(a)

ドライブレター	FCインタフェース番号	記憶領域番号	移動中フラグ
C	3714	3412	0

115
N

図6(b)

ドライブレター	FCインタフェース番号	記憶領域番号	移動中フラグ
C	3714	3412	0
D	3711	3411	0

115
N

図6(c)

ドライブレター	FCインタフェース番号	記憶領域番号	移動中フラグ
C	3714	3412	1
D	3711	3411	0

115
N

図6(d)

ドライブレター	FCインタフェース番号	記憶領域番号	移動中フラグ
C	3727	3429	0
D	3711	3411	0

【図 7】

図 7

図 7 (a)

ABH: AccessBandwidthHint

SCH: StorageCostHint

DAH: DataAvailabilityHint

384
N

記憶領域番号	FC I/F 番号	容量	帯域	記憶領域のヒント情報		
				ABH	SCH	DAH

図 7 (b)

384
N

記憶領域番号	FC I/F 番号	容量	帯域	記憶領域のヒント情報		
				ABH	SCH	DAH
3429	3727	200GB	10Gbps	10	0	10

【図 8】

図 8

800
~

図 8(a)

記憶領域作成のための設定画面

計算機	100	810	
容量	200	812	
記憶装置	301	818	
ヒント	AccessBandwidthHint	10	820
	StorageCostHint	0	822
	DataAvailabilityHint	10	824
作成		830	

801
~

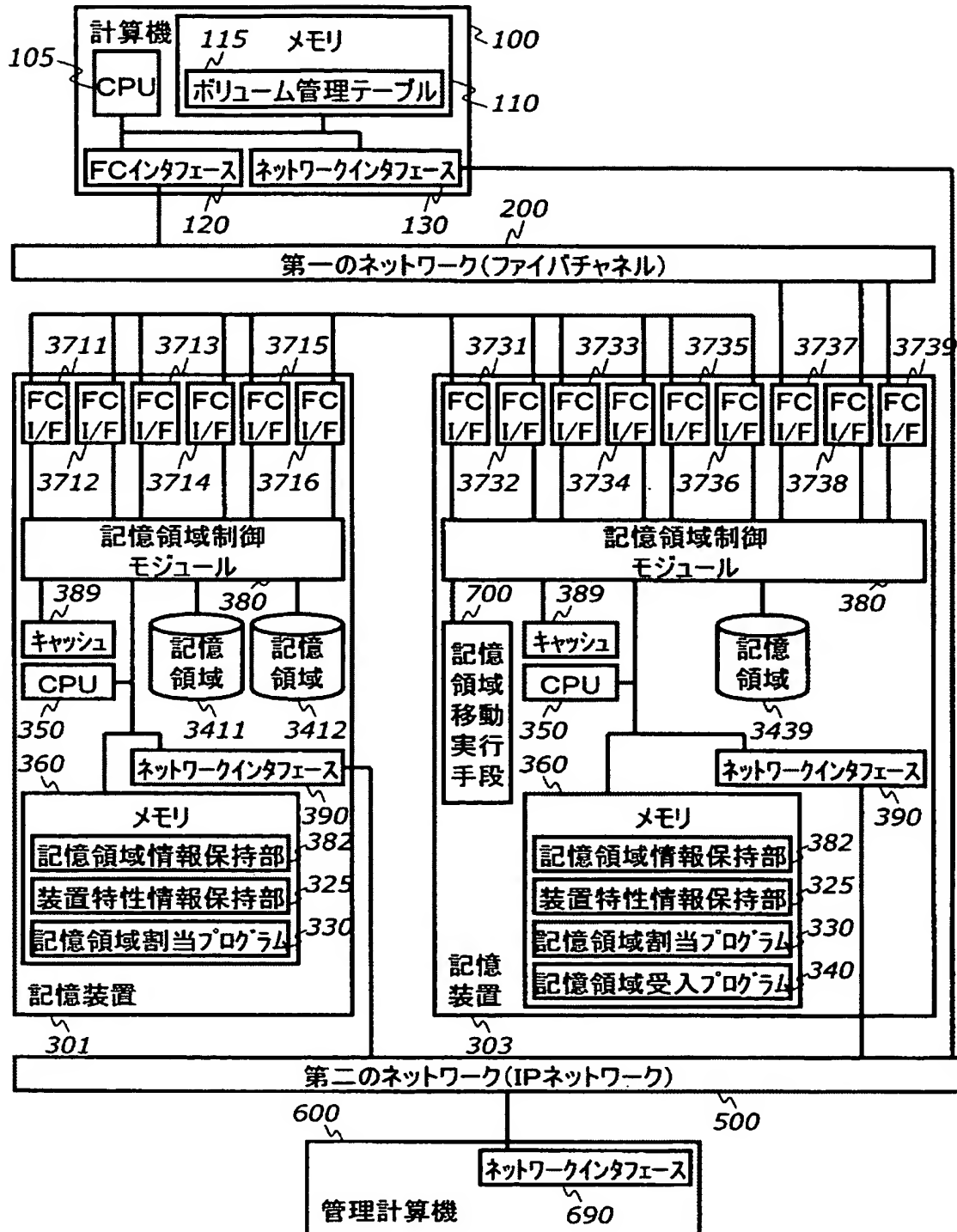
図 8(b)

記憶領域作成のための設定画面

計算機	100	810	
容量	200	812	
ボリュームポリシー	アクセス帯域・高可用重視 ▼	816	
	アクセス帯域・高可用重視 低コスト重視 アクセス帯域のみ重視		
記憶装置	801	818	
ヒント	AccessBandwidthHint	10	820
	StorageCostHint	0	822
	DataAvailabilityHint	10	824
作成		830	

【図9】

図9



【図 10】

図10

ABH: AccessBandwidthHint

SCH: StorageCostHint

DAH: DataAvailabilityHint

図10(a)

385
N

記憶領域番号	FC I/F 番号	容量	帯域	記憶領域のヒント情報			外部記憶装置	
				ABH	SCH	DAH	FC I/F 番号	記憶領域番号
3432	3738	200GB	2Gbps	10	0	10	3714	3412
3431	3737	200GB	1Gbps	1	10	0	3711	3411

図10(b)

385
N

記憶領域番号	FC I/F 番号	容量	帯域	記憶領域のヒント情報			外部記憶装置	
				ABH	SCH	DAH	FC I/F 番号	記憶領域番号
3432	3738	200GB	2Gbps	10	0	10	3714	3412
3431	3737	200GB	1Gbps	1	10	0	3711	3411
3439	0	200GB	10Gbps	10	0	10	0	0

図10(c)

385
N

記憶領域番号	FC I/F 番号	容量	帯域	記憶領域のヒント情報			外部記憶装置	
				ABH	SCH	DAH	FC I/F 番号	記憶領域番号
3431	3737	200GB	1Gbps	1	10	0	3711	3411
3432	3738	200GB	10Gbps	10	0	10	0	0

【図 11】

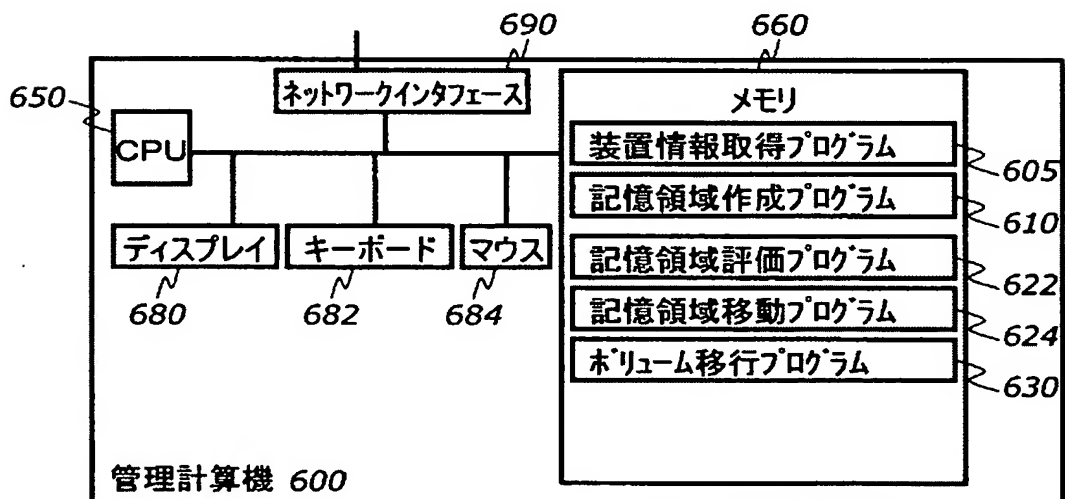
図 11

115

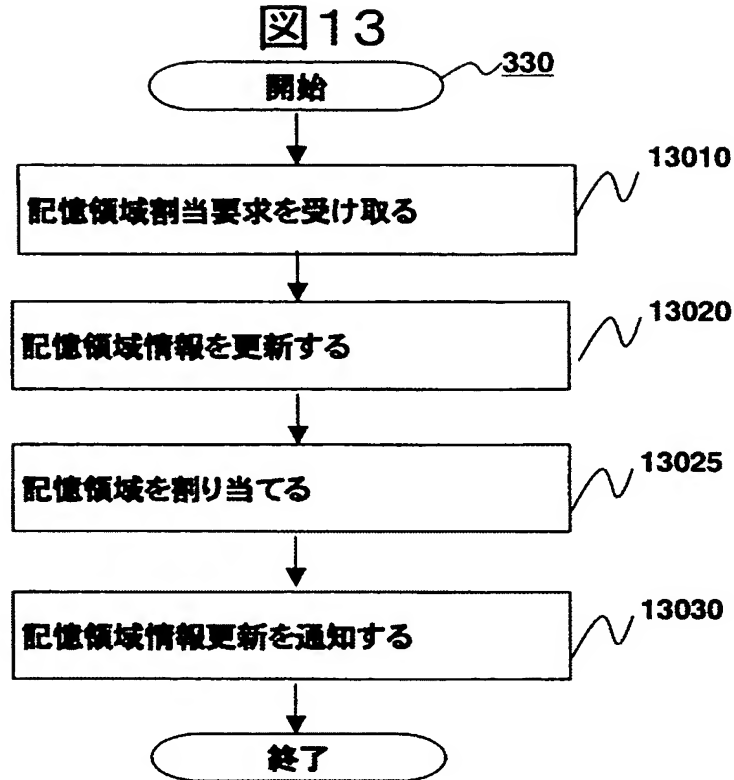
ドライブレター	FCインタフェース番号	記憶領域番号	移動中フラグ
C	3738	3432	0
D	3737	3431	0

【図 12】

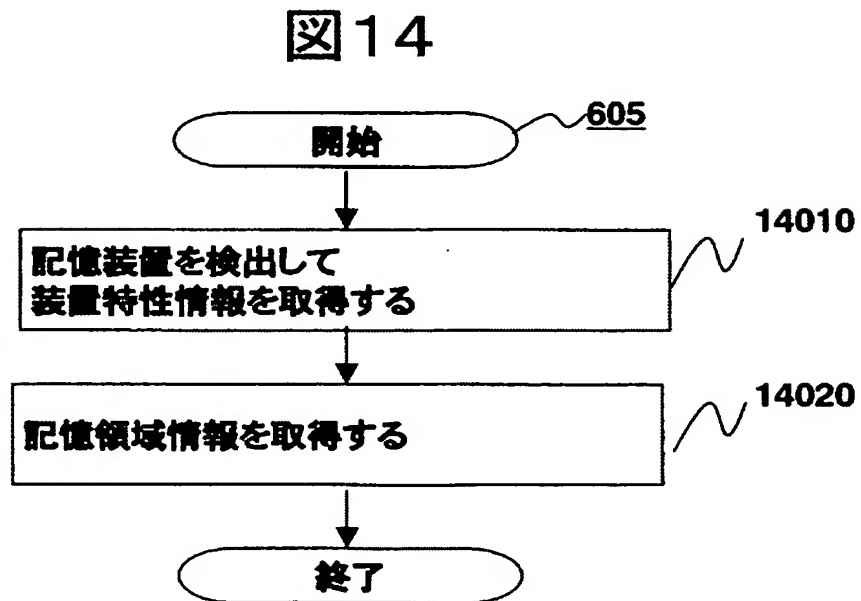
図 12



【図 13】

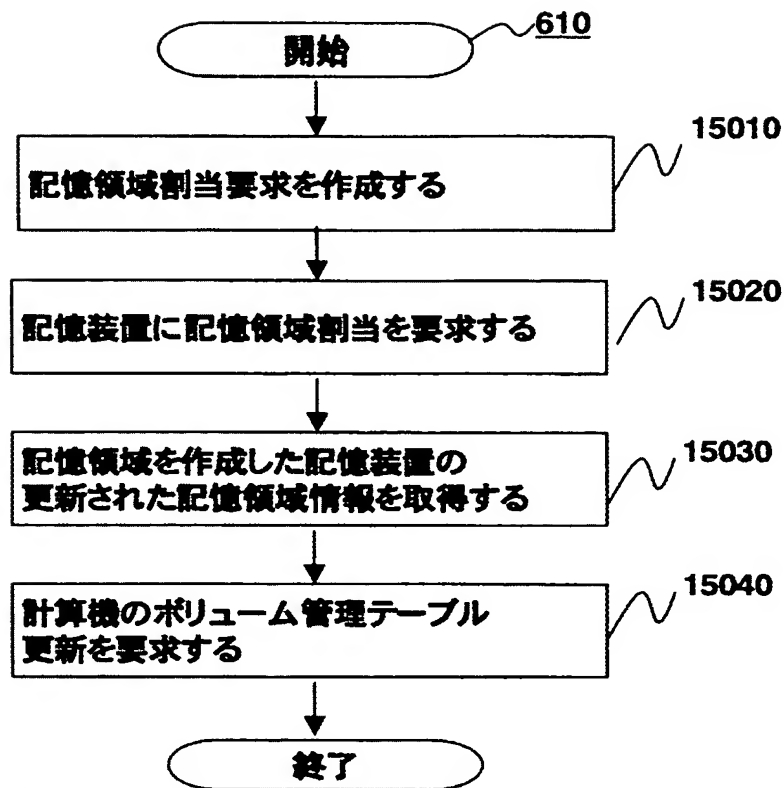


【図 14】

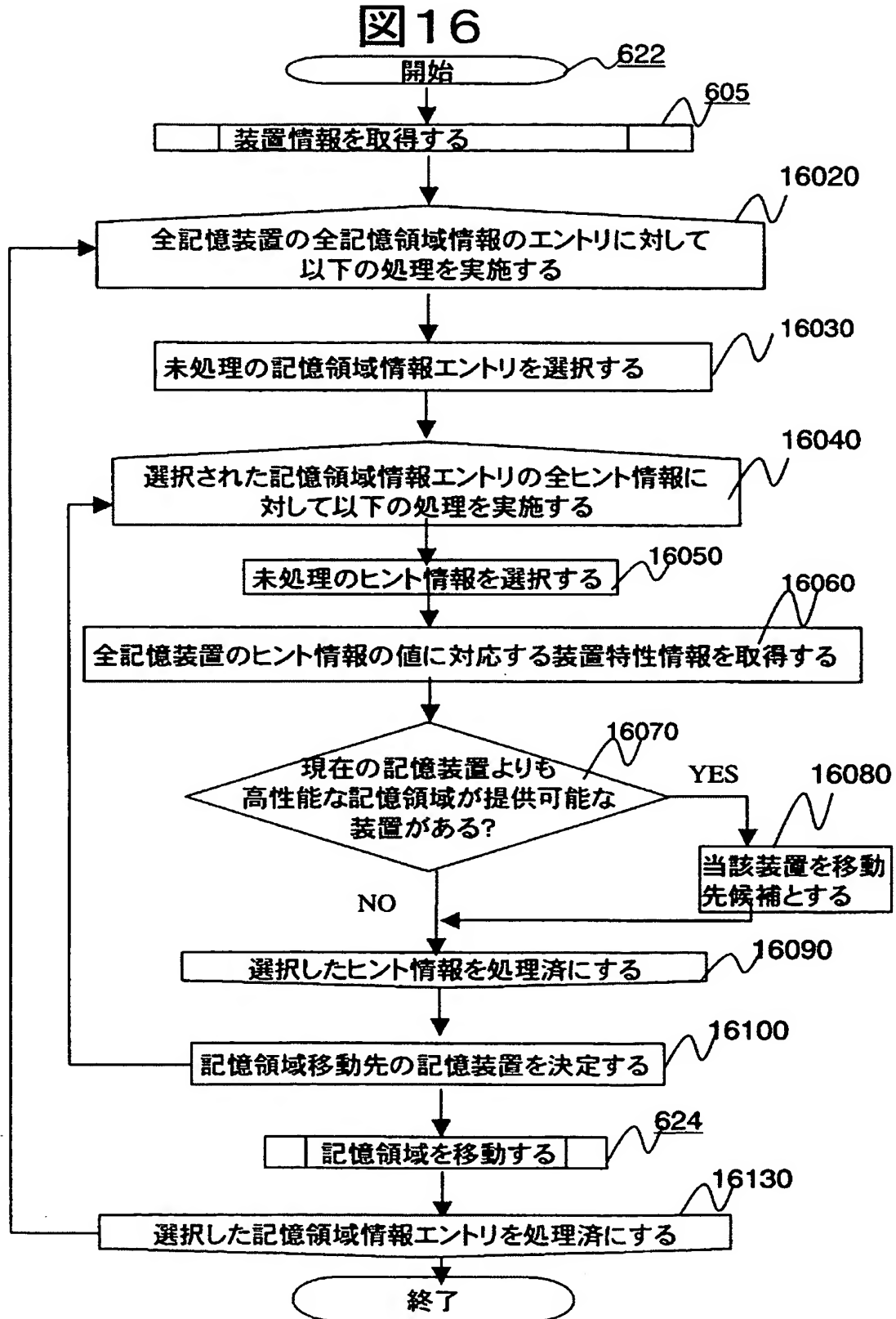


【図 15】

図 15

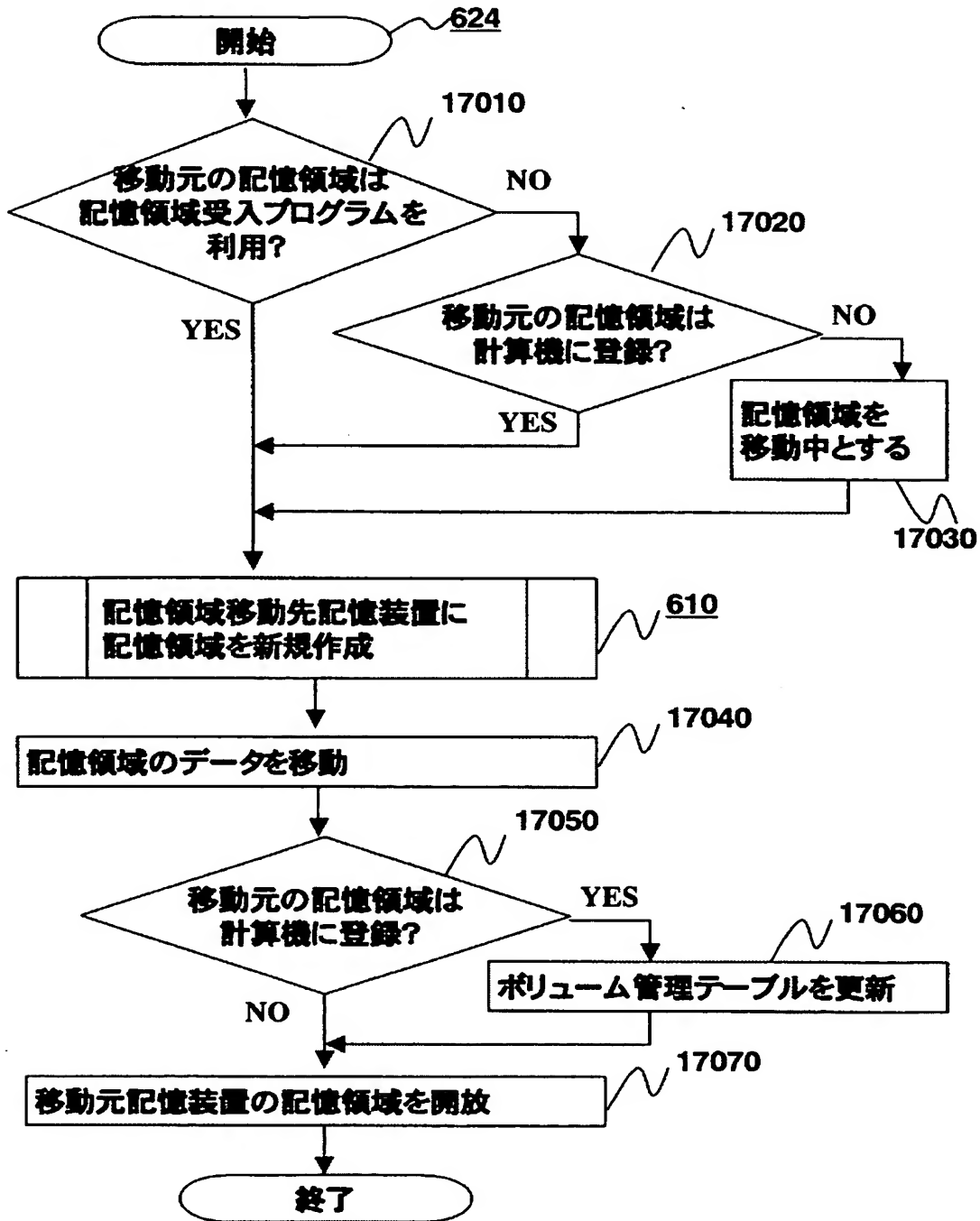


【図 16】



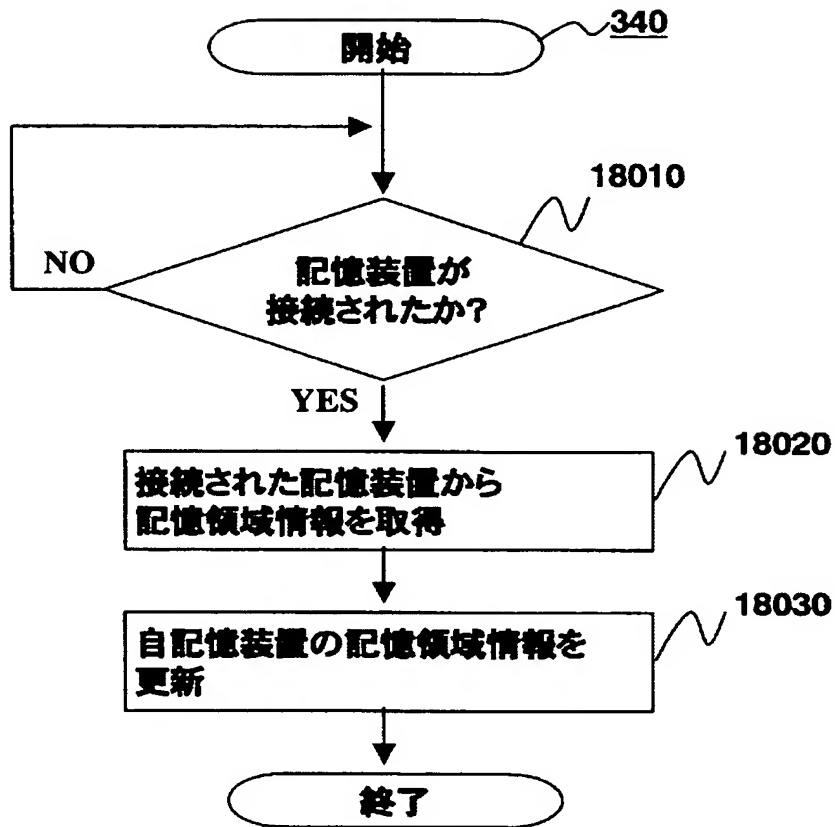
【図 17】

図 17



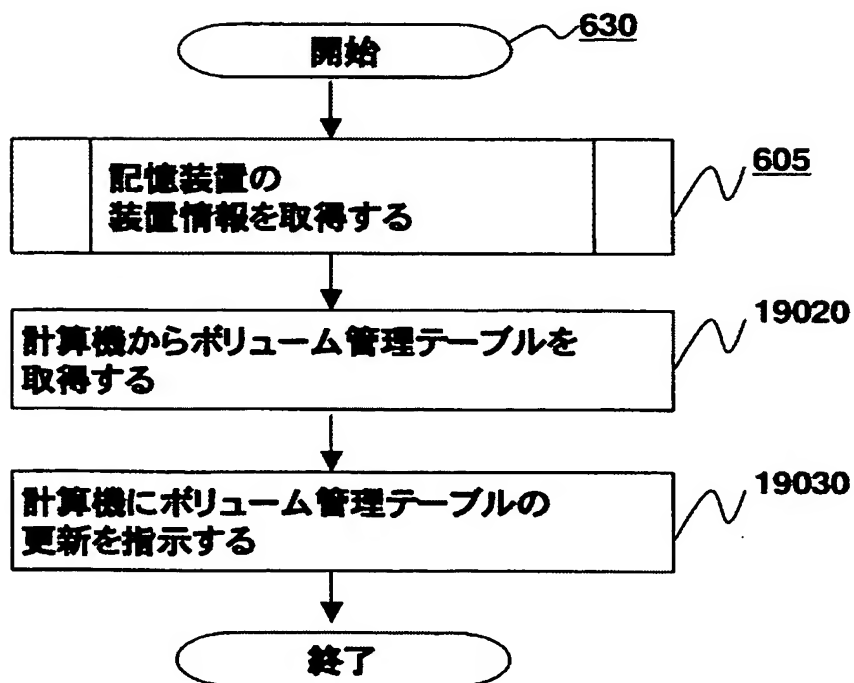
【図 18】

図 18



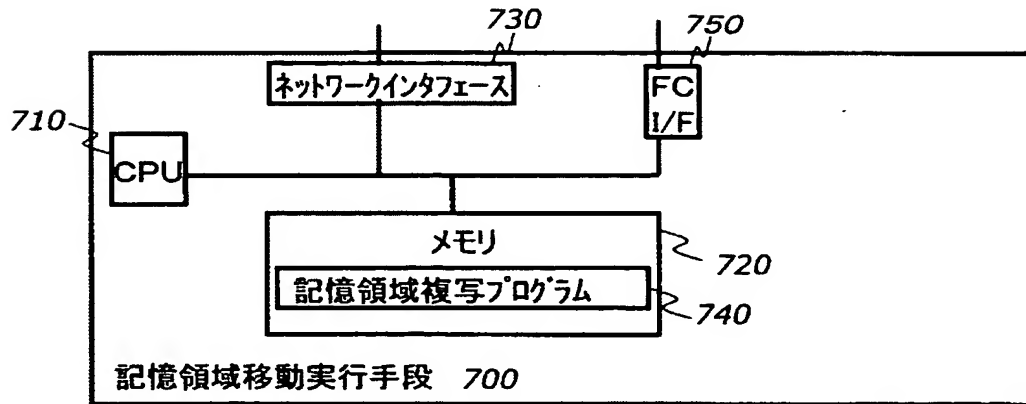
【図 19】

図 19



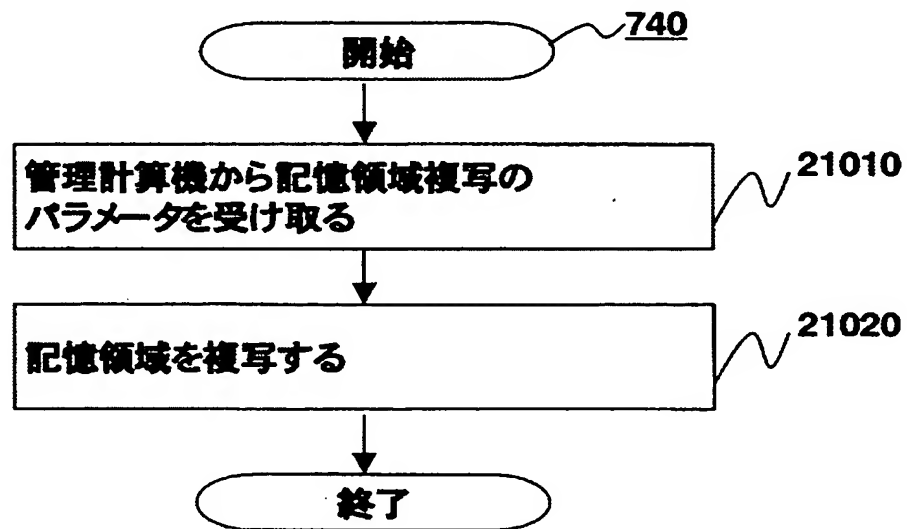
【図20】

図20



【図21】

図21



【書類名】 要約書

【要約】

【課題】

SANなどの複数の記憶装置を有する計算機システムにおいて、記憶装置の追加やSANに属する記憶装置の構成変更を検出して、既存する記憶領域を、記憶領域作成の際に与えられる「ヒント」に基づいて自動的に再配置することができなかった。

【解決手段】

計算機とネットワークを介して接続し計算機が利用するデータを格納する記憶領域を有する前記記憶装置にネットワークを介して接続する管理計算機は、記憶領域の所定の性能を示すレベルと前記記憶装置の装置特性情報との対応付けを保持し、第1の記憶装置が前記計算機に割り当てた記憶領域の性能を示すレベルを第1の記憶装置から取得する。そして、取得した記憶領域の性能を示すレベルに対応する前記第1の装置特性情報と、他の記憶装置の装置特性情報とをそれぞれ参照し、記憶装置間での前記記憶領域の性能を比較する。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願 2 0 0 4 - 0 6 6 5 4 8
受付番号	5 0 4 0 0 3 9 0 7 4 5
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 6 年 3 月 1 5 日

< 認定情報・付加情報 >

【提出日】	平成16年 3月10日
-------	-------------

特願 2 0 0 4 - 0 6 6 5 4 8

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所